

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

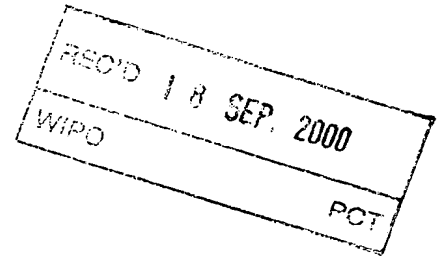
**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

**THIS PAGE BLANK (USPTO)**

BUNDESREPUBLIK DEUTSCHLAND

10/069353

**PRIORITY  
DOCUMENT**SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH RULE 17.1(a) OR (b)**Prioritätsbescheinigung über die Einreichung  
einer Patentanmeldung**

Ep 00108013

<b>Aktenzeichen:</b>	199 40 596.4	
<b>Anmeldetag:</b>	27. August 1999	
<b>Anmelder/Inhaber:</b>	Bayer AG, Leverkusen/DE	
<b>Bezeichnung:</b>	Nucleinsäuren, die für Enzymaktivitäten der Spinosyn-Biosynthese codieren	
<b>IPC:</b>	C 07 K, C 07 H, C 12 N	

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Anmeldung.

München, den 08. Juni 2000  
Deutsches Patent- und Markenamt  
Der Präsident  
Im Auftrag

### Nucleinsäuren, die für Enzymaktivitäten der Spinosyn-Biosynthese codieren

Die vorliegende Erfindung betrifft Nucleinsäuren, die für Enzymaktivitäten der Spinosyn-Biosynthese codieren, sowie die entsprechenden Enzyme per se.

5

10

Spinosyne stellen eine neue Gruppe von makrolidischen Verbindungen dar, die aus dem Actinomyceten *Saccharopolyspora spinosa* isoliert worden sind (Mertz und Yao, 1990). Sie werden zur Bekämpfung von Insekten eingesetzt (WO 97/00265, WO 94/20518, WO 93/09126, US 5670364, US 5362634, US 5227295, US 5202242). Spinosyne zeigen eine starke insektizide, jedoch keine antibakterielle Aktivität, wodurch sie von den konventionellen Makroliden, wie Tylosin, Spiramycin und Erythromycin, die keine insektizide, jedoch antimikrobielle Wirksamkeit aufweisen, unterscheidbar sind.

15

20

Die Struktur der Spinosyne setzt sich zusammen aus einem tetracyclischen Polyketidgrundgerüst (Aglycon) mit einem 12-gliedrigen Makrolidring und einem 5,6,5-cis-anti-trans-Tricyclus, sowie einem D-Forosamin- und einem 2,3,4-Tri-O-Methyl-L-Rhamnose-Zuckeranteil (Kirst et al., 1991). Mehr als 20 verschiedene natürliche Spinosyn-Derivate, der sogenannte A83543 Komplex, ist bisher beschrieben worden (WO 97/00265, WO 94/20518, WO 93/09126). Diese Derivate variieren in der Substitution von einer oder einigen Methylgruppen am tetracyclischen Grundgerüst, am Forosamin- oder am Tri-Methyl-Rhamnose-Zuckeranteil. Ein 17-Pseudoaglycon, dem der Forosamin-Zuckeranteil fehlt, ist ebenfalls aus Kulturbrühen von *S. spinosa* isoliert worden.

25

30

Die Hauptkomponenten des von *S. spinosa* gebildeten A83543 Komplexes stellen die Varianten Spinosyn A und Spinosyn D dar, die die wesentlichen Bestandteile des Produktes Spinosad darstellen (vgl. Pesticide Manual, British Crop Protection Council, 11<sup>th</sup> Ed., 1997, Seite 1272 und Dow Elanco trade magazine Down to Earth, Vol. 52, NO.: 1, 1997 und die darin zitierte Literatur).



Aufbauend auf Untersuchungen zum Einbau von  $C^{13}$ -markiertem Acetat, Propionat, Butyrat oder Isobutytrat konnte gezeigt werden, dass die Biosynthese von A83543 einem Polyketid-Biosyntheseweg folgt (Nakatsukasa et al., 1990). Polyketide werden durch multifunktionelle Enzyme, den sog. Polyketidsynthasen (PKS's) aus kurz-

5 kettigen Säurebausteinen wie Acetat, Propionat oder Butyrat aufgebaut. Ähnlich wie die verwandten Fettsäuresynthasen (FAS's) katalysieren sie decarboxylierende Polykondensationsschritte der als CoA-Thioester aktivierten Bausteine. Während FAS's nach jedem Kondensationsschritt eine vollständige Reduktion der intermediär an der wachsenden Polyketidkette entstehenden  $\beta$ -Oxoester durch Ketoreduktion,

10 Dehydratation und Enoylreduktion katalysieren, können PKS's bestimmte Reduktionsschritte auslassen. Modulare Typ I PKS's bestehen aus einem oder mehreren großen multifunktionalen Proteinen. Iterative Typ II PKS's stellen dagegen einen Komplex aus weitgehend monofunktionalen Proteinen dar.

15 Die enzymatischen Aktivitäten von modularen Typ I PKS's lassen sich zu sogenannten Modulen zusammenfassen. Hierbei trägt ein Modul eine Anordnung von drei enzymkatalytisch aktiven Domänen, die zu einer Verlängerung der wachsenden Polyketidkette um eine biosynthetische Verlängerungseinheit führen. Bei diesen Domänen handelt es sich um eine  $\beta$ -Ketoacyl:Acyl Carrier Protein

20 Synthase-Domäne, eine Acyltransferase-Domäne und eine  $\beta$ -Ketoacyl:Acyl Carrier Protein-Domäne. Ein Modul kann auch eine Ketoreduktase-, eine Dehydratase-, eine Enoylreduktase- und eine Thioesterase-Domäne tragen. Ein sog. Ladungsmodul, das am Beginn der Biosynthese steht kann von den genannten Domänen lediglich eine Acyltransferase-Domäne und eine  $\beta$ -Ketoacyl:Acyl-Carrier Protein-Domäne tragen,

25 sowie eine enzymatisch inaktive  $\beta$ -Ketoacyl:Acyl Carrier Protein Synthase-Domäne. Eine Polyketidsynthase-Domäne umfasst jeweils eine dieser genannten enzymatischen Aktivitäten.

---

Aufgrund der potenten insektiziden Wirkung sowie der bemerkenswerten Struktur

30 der Spinosyne besteht ein großes Interesse, die genetischen Informationen für deren Biosynthese zu entschlüsseln.

Gegenstand der Erfindung sind Nucleinsäuren, welche zumindest eine Region umfassen, die für eine Enzymaktivität codiert, welche an der Biosynthese von Spinosynen beteiligt ist.

5

Die vorliegende Erfindung stellt ein Cluster von offenen Leserahmen (ORF's) bereit, deren Translationsprodukte an der Biosynthese von Spinosynen beteiligt sind. Weiterhin werden zusätzliche Gene bzw. ORF's bereitgestellt, die außerhalb des ca. 120 kb großen Spinosyn-Biosyntheseclusters liegen, und deren Translationsprodukte an der Rhamnose-Zuckerbiosynthese beteiligt sind.

10

Bei den erfindungsgemäßen Nucleinsäuren handelt es sich insbesondere um einzelsträngige oder doppelsträngige Desoxyribonucleinsäuren (DNA) oder Ribonucleinsäuren (RNA). Bevorzugte Ausführungsformen sind Fragmente genomischer DNA und cDNA's.

15

Der Ausdruck "zumindest eine Region", wie er hierin verwendet wird, bedeutet, dass die erfindungsgemäße Nucleinsäure eine oder mehrere Sequenzen umfassen kann, welche jeweils für einzelne Aktivitäten codieren, die Schritte bei der Synthese von Spinosynen durchführen. Es werden demnach auch Nucleinsäuren als erfindungsgemäß betrachtet, die nur für eine einzige Enzymaktivität der Spinosyn-Biosynthese codieren.

20

Der Ausdruck "Enzymaktivität", wie er hierin verwendet wird, bedeutet, dass ausgehend von den hierin betrachteten Nucleinsäuren zumindest derjenige Teil eines vollständigen Enzyms exprimiert werden kann, der noch die Katalyseeigenschaften des Enzyms ausübt.

25

Insbesondere codieren die erfindungsgemäßen Nucleinsäuren für Enzymaktivitäten von Polyketidsynthasen, Methyltransferasen, Epimerasen, Glycosyltransferasen,

30

Aminotransferasen, Dimethyltransferasen, Reduktasen, Dehydratasen und/oder Cyclisierungsenzymen.

5 Bevorzugt handelt es sich bei den erfindungsgemäßen Nucleinsäuren um DNA-Fragmente, die genomischer DNA von *S. spinosa* entsprechen.

Besonders bevorzugt umfassen die erfindungsgemäßen Nucleinsäuren zumindest eine Sequenz ausgewählt aus

- 10
- (a) den Sequenzen gemäß SEQ ID NOS: 1, 2, 3, 4, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 52 oder 54,
- 15
- (b) zumindest 14 Basenpaare langen Teilsequenzen der unter (a) definierten Sequenzen,
- (c) Sequenzen, welche an die unter (a) definierten Sequenzen hybridisieren,
- 20
- (d) Sequenzen, welche eine zumindest 70 %ige, bevorzugt eine 80 %ige, besonders bevorzugt eine 90 %ige Identität zu den unter (a) definierten Sequenzen aufweisen,
- 25
- (e) Sequenzen, welche zu den unter (a) definierten Sequenzen komplementär sind, und
- (f) Sequenzen, welche aufgrund der Degeneriertheit des genetischen Codes für dieselbe Aminosäuresequenz kodieren wie die unter (a) bis (d) definierten Sequenzen.
-

Der Ausdruck "hybridisieren", wie er hierin verwendet wird, beschreibt den Vorgang, bei welchem ein einzelsträngiges Nucleinsäuremolekül mit einem komplementären Strang eine Basenpaarung eingeht. Auf diese Weise können beispielsweise ausgehend von genomischer DNA aus Organismen, die phylogenetisch mit *S. spinosa* verwandt sind und die Fähigkeit der Biosynthese von Spinosynen besitzen, DNA-Fragmente isoliert werden, welche dieselben Eigenschaften wie die aus *S. spinosa* isolierten Fragmente aufweisen.

Bevorzugte Hybridisierungsbedingungen sind nachstehend angegeben: Hybridisierungslösung: 5 x SSC; Blocking Reagents (Roche Diagnostics GmbH, Mannheim, Deutschland), 1 %; N-Lauroylsarcosin, 0.1 %; SDS (Sodiumdodecylsulfate) 0.02 %; Hybridisierungstemperatur: 60°C; erster Waschschrift: 2 x SSC bei 60°C; zweiter Waschschrift: 2 x SSC bei 60°C; bevorzugt zweiter Waschschrift: 0.5 x SSC bei 60°C; besonders bevorzugt zweiter Waschschrift: 0.2 x SSC bei 60°C.

Der Grad der Identität der Nucleinsäuren wird vorzugsweise bestimmt mit Hilfe des Programms GAP aus dem Programmpaket GCG (Devereux et al., 1984), Version 9.1 unter Standardeinstellungen.

Besonders hervorgehoben werden Nucleinsäuren, die

- (1) entweder alle Sequenzen, die für Schritte der Forosamin- und Trimethyl-Rhamnose-Biosynthese codieren, umfassen, insbesondere die Sequenzen gemäß SEQ ID NOS: 4 und 51, oder
- (2) alle Sequenzen, die für Schritte der Polyketidsynthese codieren, umfassen, insbesondere die Sequenzen gemäß SEQ ID NOS: 5 und 6, oder
- (3) alle Sequenzen, die für alle Schritte der Forosamin-, Trimethyl-Rhamnose- und Polyketidsynthese codieren, umfassen, insbesondere die Sequenzen gemäß SEQ ID NOS: 1, 2, 3 und 51.

5

Alle zur Spinosyn-Biosynthese oder zur Synthese von Vorstufen, wie sie nachstehend definiert sind, benötigten DNA-Sequenzen können sich somit auf einem einzelnen Vektor befinden. Diese Nucleinsäuren können aber auch auf zwei oder mehreren Vektoren vorliegen und gleichzeitig oder nacheinander in einer Wirtszelle exprimiert werden.

Alle ORF's der erfindungsgemäßen Nucleinsäuren können von ihren eigenen Promotoren oder von heterologen Promotoren angeschaltet werden.

10

Gegenstand der vorliegenden Erfindung sind auch die regulatorischen Regionen, welche natürlicherweise, d.h. im Ursprungsorganismus *S. spinosa*, die Transkription der erfindungsgemäßen Nucleinsäuren kontrollieren.

15

Der Ausdruck "regulatorische Regionen", wie er hierin verwendet wird, bezieht sich auf Promotoren, Repressor- oder Aktivator-Bindungsstellen, Repressor- oder Aktivatorsequenzen, und Terminatoren. Ferner sind genetisch mobile Elemente, welche natürlicherweise, d.h. im Ursprungsorganismus *S. spinosa* vorkommen, ebenfalls von diesem Ausdruck umfasst. Solche genetisch mobilen Elemente können transposable oder mobilisierbare Elemente oder funktionelle Teile davon, IS-Elemente oder andere Insertionselemente sein. Weiterhin sind auch amplifizierbare DNA-Elemente (Amplifiable Units of DNA, AUD; Fishman and Hershberger, 1983), welche natürlicherweise, d.h. im Ursprungsorganismus *S. spinosa* vorkommen, von diesem Ausdruck umfasst. Die Erfindung betrifft auch jede Kombination dieser regulatorischen Regionen untereinander oder mit heterologen DNA-Fragmenten, wie z.B. Promotoren, Repressor oder Aktivator-Bindungsstellen, transposablen, mobilisierbaren oder transduzierbaren Elementen.

20

25

30

---

Gegenstand der vorliegenden Erfindung sind weiterhin DNA-Konstrukte, die zumindest eine erfindungsgemäße Nucleinsäure und einen heterologen Promotor umfassen.

Der Ausdruck "heterologer Promotor", wie er hierin verwendet wird, bezieht sich auf einen Promotor, der im Ursprungsorganismus nicht die Expression des betreffenden Gens (ORF's) kontrolliert.

5 Die Auswahl von heterologen Promotoren ist davon abhängig, ob zur Expression pro- oder eukaryotische Zellen oder zellfreie Systeme verwendet werden. Ein bevorzugtes Beispiele für einen heterologen Promotor ist der Promotor des *mel*-Gens aus dem Vektor pIJ702 (The John Innes Foundation, Norwich, UK 1985). Die heterologe Expression kann z.B. eingesetzt werden, um zu einer Steigerung der Produktion von Spinosyn im Vergleich zum natürlichen Spinosyn-Produzenten zu gelangen.

10 Gegenstand der Erfindung sind ferner Vektoren, die zumindest eine der erfindungsgemäßen Nucleinsäuren enthalten. Als Vektoren können alle in 15 molekularbiologischen Laboratorien verwendeten Phagen, Plasmide, Phagmide, Phasmide, Cosmide, YACs, BACs, künstliche Chromosomen oder Partikel, die für einen Partikelbeschluss geeignet sind, verwendet werden.

20 Bevorzugte Vektoren sind solche, die geeignet sind zur Übertragung und Replikation in oder Integration ins Genom von Actinomyceten, wie z.B. Streptomyceten.

25 Besonders bevorzugt sind BAC-Vektoren, insbesondere solche BAC-Vektoren, die in der Weise modifiziert sind, dass sie zur Übertragung in Actinomyceten, wie z.B. Streptomyceten, sowie zur Replikation in, oder Integration ins Genom von Actinomyceten, wie z.B. Streptomyceten, geeignet sind.

BAC-Vektoren (Bacterial Artificial Chromosome) sind entwickelt worden zur Klonierung von großen DNA-Fragmenten (Shizuya et al., 1992). Es handelt sich um "single-copy" Plasmide mit einem F-Faktor Origin, die DNA-Fragmente mit einer 30 durchschnittlichen Größe von 120 Kilobasenpaaren (kb) tragen können. Sie sind replizierbar in *Escherichia coli*. Der BAC-Vektor pBeloBAC11 (Kim et al., 1996)

trägt einen T7 und einen SP6 Promotor, welche die Klonierungsstelle flankieren und als Startbereich für Sequenzierungsprimer sowie zur Generierung von RNA-Transkripten verwendet werden können.

5 Am meisten bevorzugt sind die am 20. August 1999 bei der Deutschen Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ), Mascheroder Weg 1b, D-38124 Braunschweig, in Übereinstimmung mit den Bestimmungen des Budapester Vertrages unter den Hinterlegungsnummern DSM 13010, DSM 13011 und DSM 13012 hinterlegten BAC-Shuttleklone, die Gegenstand der vorliegenden Erfindung sind.

10

Die hinterlegten BAC-Shuttleklone P11/G6, P8/G11 und P11/B10 tragen jeweils ein mindestens 100 kb großes DNA-Fragment aus *S. spinosa*. Die Klonen P11/G6 und P11/B10 tragen jeweils einen Teil der Nucleinsäuresequenz gemäß SEQ ID NO: 4, 15 sowie die angrenzenden vollständigen Nucleinsäuresequenzen gemäß SEQ ID NOS: 5 und 6, sowie einen an die Nucleotidsequenz gemäß SEQ ID NO: 6 3'-angrenzende DNA-Bereich (Abb. 7). Der Klon P8/G11 trägt einen Teil der Nucleinsäuresequenz gemäß SEQ ID NO: 6, die vollständigen Nucleinsäuresequenzen gemäß SEQ ID NOS: 5 und 4, sowie einen an die Sequenz gemäß SEQ ID NO: 4 3'-angrenzenden 20 DNA-Bereich (Abb. 7).

20

Gegenstand der vorliegenden Erfindung sind auch Wirtszellen, die zumindest eine der erfindungsgemäßen Nucleinsäuren enthalten. Als Wirtszelle eignen sich sowohl prokaryotische Zellen, vorzugsweise Actinomyceten, besonders bevorzugt Strepto- 25 myceten, als auch eukaryotische Zellen, wie Säugerzellen, Pflanzenzellen oder Hefezellen.

25

In besonderer Weise können die erfindungsgemäßen Nucleinsäuren in pflanzliche Zellen übertragen und exprimiert werden. Hierdurch können transgene Pflanzen 30 hergestellt werden, die das pflanzenschützende, insektizide Spinosyn bzw. Derivate davon produzieren. Eine Übertragung der erfindungsgemäßen Nucleinsäuren in die

30

Pflanzenzellen oder pflanzliche Zellkulturen kann mit üblichen Verfahren u.a. auch durch Partikelbeschuss erfolgen.

5 Gegenstand der vorliegenden Erfindung sind weiterhin die Polypeptide, die von den erfindungsgemäßen Nucleinsäuren codiert werden. Die erfindungsgemäßen Polypeptide können ein vollständiges Enzym darstellen, das einen Schritt der Spinosyn-Biosynthese katalysiert. Jedoch sind auch solche Polypeptide von der Erfindung erfasst, die nur einen Teil der vollständigen Aminosäuresequenz des betreffenden Enzyms aufweisen.

10

Der Ausdruck "Teilsequenz", wie er hierin verwendet wird, bezieht sich somit auf die Aminosäuresequenz eines Polypeptids, das noch die Aktivität des entsprechenden vollständigen Enzyms oder einer enzymatisch aktiven Domäne ausüben kann.

15

Im Folgenden werden bevorzugte erfindungsgemäße Nucleinsäuren und Polypeptide mit Bezug auf die entsprechenden SEQ ID NOS näher charakterisiert.

SEQ ID NOS: 7 und 8, ORF1:

20

Nucleotidpositon 828 bis 1 der SEQ ID NO: 4, 275 Aminosäuren;  
das ableitbare Genprodukt ist eine Methyltransferase.

SEQ ID NOS: 9 und 10, ORF2:

25

Nucleotidposition 1.283 bis 2.455 der SEQ ID NO: 4, 390 Aminosäuren;  
das ableitbare Genprodukt ist eine Glycosyltransferase.

SEQ ID NOS: 11 und 12, ORF3:

Nucleotidposition 2.495 bis 3.247 der SEQ ID NO: 4, 250 Aminosäuren;  
das ableitbare Genprodukt ist eine Methyltransferase.



SEQ ID NOS: 13 und 14, ORF4:

Nucleotidposition 4.440 bis 3.253 der SEQ ID NO: 4, 395 Aminosäuren;  
das ableitbare Genprodukt ist eine Methyltransferase.

5 SEQ ID NOS: 15 und 16, ORF5:

Nucleotidposition 4.578 bis 6.197 der SEQ ID NO: 4, 539 Aminosäuren;  
das ableitbare Genprodukt ist ein C-C verknüpfendes Enzym, das Cyclisierungsreaktionen durchführt.

10 SEQ ID NOS: 17 und 18, ORF6:

Nucleotidposition 6.211 bis 7.404 der SEQ ID NO: 4, 397 Aminosäuren;  
das ableitbare Genprodukt ist eine Methyltransferase.

SEQ ID NOS: 19 und 20, ORF7:

15 Nucleotidposition 7.401 bis 8.300 der SEQ ID NO: 4, 299 Aminosäuren;  
das ableitbare Genprodukt ist eine Methyltransferase.

SEQ ID NOS: 21 und 22, ORF8:

20 Nucleotidposition 8.300 bis 9.466 der SEQ ID NO: 4, 388 Aminosäuren;  
das ableitbare Genprodukt ist ein Enzym, das an Cyclisierungsreaktionen beteiligt ist.

SEQ ID NOS: 23 und 24, ORF9:

Nucleotidposition 10.572 bis 9.562 der SEQ ID NO: 4, 336 Aminosäuren.  
das ableitbare Genprodukt ist eine 2,3-Reduktase.

25

SEQ ID NOS: 25 und 26, ORF10:

Nucleotidposition 12.029 bis 10.569 der SEQ ID NO: 4, 486 Aminosäuren;  
das ableitbare Genprodukt ist eine 2,3-Dehydratase.

---

SEQ ID NOS: 27 und 28, ORF11:

Nucleotidposition 12.549 bis 12.109 der SEQ ID NO: 4, 146 Aminosäuren; das ableitbare Genprodukt hat Homologien zu einer Thioesterase.

5 SEQ ID NOS: 29 und 30, ORF12:

Nucleotidposition 13.865 bis 12.546 der SEQ ID NO: 4, 439 Aminosäuren; das ableitbare Genprodukt ist eine Glykosyltransferase.

SEQ ID NOS: 31 und 32, ORF13:

10 Nucleotidposition 14.245 bis 15.633 der SEQ ID NO: 4, 462 Aminosäuren; das ableitbare Genprodukt ist eine 3,4-Dehydratase.

SEQ ID NOS: 33 und 34, ORF14:

15 Nucleotidposition 15.671 bis 16828 der SEQ ID NO: 4, 385 Aminosäuren; das ableitbare Genprodukt ist eine 4-Aminotransferase.

SEQ ID NOS: 35 und 36, ORF15:

Nucleotidposition 16.831 bis 17.580 der SEQ ID NO: 4, 249 Aminosäuren; das ableitbare Genprodukt ist eine N-Dimethyltransferase.

20 SEQ ID NOS: 37 und 38, ORF16:

Nucleotidposition 18.930 bis 18.205 der SEQ ID NO: 4, 241 Aminosäuren; das ableitbare Genprodukt ist eine 3,4-Reduktase.

25 SEQ ID NOS: 39 und 40, ORF17:

Nucleotidposition 19.025-19.861 der SEQ ID NO: 4, 278 Aminosäuren; das ableitbare Genprodukt ist ein Transkriptions-Regulator.

---

SEQ ID NOS: 41 und 42, ORF18:

Nucleotidpositionen 116-7903 der SEQ ID NO: 5, Aminosäurepositionen 1 bis 2595:

Nucleotidpositionen 128-1402, Aminosäure Positionen 5-429 codieren eine  $\beta$ -

5 Ketoacyl:Acyl Carrier Protein Synthase-Domäne;

Nucleotidpositionen 1691-2656, Aminosäurepositionen 526-847 codieren eine Acyltransferase-Domäne;

Nucleotidpositionen 2798-3052, Aminosäurepositionen 895-979 codieren eine  $\beta$ -Ketoacyl:Acyl Carrier Protein-Domäne;

10 Nucleotidpositionen 3107-4372, Aminosäurepositionen 998-1419 codieren eine  $\beta$ -Ketoacyl:Acyl Carrier Protein Synthase-Domäne;

Nucleotidpositionen 4688-5662, Aminosäurepositionen 1525-1849 codieren eine Acyltransferase-Domäne;

15 Nucleotidpositionen 6587-7138, Aminosäurepositionen 2158-2341 codieren eine Ketoreduktase-Domäne;

Nucleotidpositionen 7409-7666, Aminosäurepositionen 2432-2517 codieren eine  $\beta$ -Ketoacyl:Acyl Carrier Protein-Domäne.

SEQ ID NOS: 43 und 44, ORF19:

20 Nucleotidpositionen 7921-14379 der SEQ ID NO: 5, Aminosäurepositionen 1 bis 2152:

Nucleotidpositionen 8029-9318, Aminosäurepositionen 37-466 codieren eine  $\beta$ -Ketoacyl:Acyl Carrier Protein Synthase-Domäne;

25 Nucleotidpositionen 9634-10608, Aminosäurepositionen 572-896 codieren eine Acyltransferase-Domäne;

Nucleotidpositionen 10705-11259, Aminosäurepositionen 929-1113 codieren eine Dehydratase-Domäne;

Nucleotidpositionen 12043-13080, Aminosäurepositionen 1375-1720 codieren eine Enoylreduktase-Domäne;

30 Nucleotidpositionen 13093-13635, Aminosäurepositionen 1725-1905 codieren eine Ketoreduktase-Domäne;

Nucleotidpositionen 13885-14142, Aminosäurepositionen 1989-2074 codieren eine  $\beta$ -Ketoacyl:Acyl Carrier Protein-Domäne.

SEQ ID NOS: 45 und 46, ORF20:

5 Nucleotidpositionen 14424-23936 der SEQ ID NO: 5, Aminosäurepositionen 1 bis 3170:

Nucleotidpositionen 14523-15824, Aminosäurepositionen 34-467 codieren eine  $\beta$ -Ketoacyl:Acyl Carrier Protein Synthase-Domäne;

10 Nucleotidpositionen 16110-17075, Aminosäurepositionen 563-884 codieren eine Acyltransferase-Domäne;

Nucleotidpositionen 17997-18536, Aminosäurepositionen 1192-1371 codieren eine Ketoreduktase-Domäne;

Nucleotidpositionen 18795-19052, Aminosäurepositionen 1458-1543 codieren eine  $\beta$ -Ketoacyl:Acyl Carrier Protein-Domäne;

15 Nucleotidpositionen 19107-20387, Aminosäurepositionen 1562-1988 codieren eine  $\beta$ -Ketoacyl:Acyl Carrier Protein Synthase-Domäne;

Nucleotidpositionen 20718-21692, Aminosäurepositionen 2099-2423 codieren eine Acyltransferase-Domäne;

20 Nucleotidpositionen 22620-23171, Aminosäurepositionen 2733-2916 codieren eine Ketoreduktase-Domäne;

Nucleotidpositionen 23436-23693, Aminosäurepositionen 3005-3090 codieren eine  $\beta$ -Ketoacyl:Acyl Carrier Protein-Domäne.

SEQ ID NOS: 47 und 48, ORF21:

25 Nucleotidpositionen 23983-38757 der SEQ ID NO: 5, Aminosäurepositionen 1 bis 4924:

~~Nucleotidpositionen 24082-25392, Aminosäurepositionen 34-470 codieren eine  $\beta$ -Ketoacyl:Acyl Carrier Protein Synthase-Domäne;~~

30 ~~Nucleotidpositionen 25696-26661, Aminosäurepositionen 572-893 codieren eine Acyltransferase-Domäne;~~

Nucleotidpositionen 26761-27315, Aminosäurepositionen 927-1111 codieren eine Dehydratase-Domäne;

Nucleotidpositionen 28231-28782, Aminosäurepositionen 1417-1600 codieren eine Ketoreduktase-Domäne;

5 Nucleotidpositionen 29035-29265, Aminosäurepositionen 1685-1761 codieren eine  $\beta$ -Ketoacyl:Acyl Carrier Protein-Domäne;

Nucleotidpositionen 29329-30624, Aminosäurepositionen 1783-2214 codieren eine  $\beta$ -Ketoacyl:Acyl Carrier Protein Synthase-Domäne;

10 Nucleotidpositionen 30928-31902, Aminosäurepositionen 2316-2640 codieren eine Acyltransferase-Domäne;

Nucleotidpositionen 32827-33378, Aminosäurepositionen 2949-3132 codieren eine Ketoreduktase-Domäne;

Nucleotidpositionen 33652-33900, Aminosäurepositionen 3224-3306 codieren eine  $\beta$ -Ketoacyl:Acyl Carrier Protein-Domäne;

15 Nucleotidpositionen 33952-35262, Aminosäurepositionen 3324-3760 codieren eine  $\beta$ -Ketoacyl:Acyl Carrier Protein Synthase-Domäne;

Nucleotidpositionen 35554-36522, Aminosäurepositionen 3858-4180 codieren eine Acyltransferase-Domäne;

20 Nucleotidpositionen 37453-37998, Aminosäurepositionen 4491-4672 codieren eine Ketoreduktase-Domäne;

Nucleotidpositionen 38254-38511, Aminosäurepositionen 4758-4843 codieren eine  $\beta$ -Ketoacyl:Acyl Carrier Protein-Domäne.

SEQ ID NOS: 49 und 50, ORF22:

25 Nucleotidpositionen 38808-50000 der SEQ ID NO: 5 und die Nukleotidpositionen 1 bis 5574 der SEQ ID NO: 6, Aminosäurepositionen 1 bis 5588:

~~Nucleotidpositionen 38907-40226 der SEQ ID NO: 5, Aminosäurepositionen 34-473~~  
codiert eine  $\beta$ -Ketoacyl:Acyl Carrier Protein Synthase-Domäne;

30 Nucleotidpositionen 40494-41453 der SEQ ID NO: 5, Aminosäurepositionen 563-882 codieren eine Acyltransferase-Domäne;

Nucleotidpositionen 41556-42119 der SEQ ID NO: 5, Aminosäurepositionen 917-1104 codieren eine Dehydratase-Domäne;

Nucleotidpositionen 43017-43568 der SEQ ID NO: 5, Aminosäurepositionen 1404-1587 codieren eine Ketoreduktase-Domäne;

5 Nucleotidpositionen 43833-44090 der SEQ ID NO: 5, Aminosäurepositionen 1676-1761 codieren eine  $\beta$ -Ketoacyl:Acyl-Carrier Protein Domäne;

Nucleotidpositionen 44151-45473 der SEQ ID NO: 5, Aminosäurepositionen 1782-2222 codieren eine  $\beta$ -Ketoacyl:Acyl Carrier Protein Synthase-Domäne;

10 Nucleotidpositionen 45765-46730 der SEQ ID NO: 5, Aminosäurepositionen 2320-2641 codieren eine Acyltransferase-Domäne;

Nucleotidpositionen 46827-47459 der SEQ ID NO: 5, Aminosäurepositionen 2674-2884 codieren eine Dehydratase-Domäne;

Nucleotidpositionen 48378-48935 der SEQ ID NO: 5, Aminosäurepositionen 3191-3376 codieren eine Ketoreduktase-Domäne;

15 Nucleotidpositionen 49182-49412 der SEQ ID NO: 5, Aminosäurepositionen 3459-3535 codieren eine  $\beta$ -Ketoacyl:Acyl Carrier Protein-Domäne;

Nucleotidpositionen 49482-50000 der SEQ ID NO: 5 und Nucleotidposition 1 bis 759 der SEQ ID NO: 6, Aminosäurepositionen 3559-3984 codieren eine  $\beta$ -Ketoacyl:Acyl Carrier Protein Synthase-Domäne;

20 Nucleotidpositionen 1084-2049 der SEQ ID NO: 6, Aminosäurepositionen 4093-4414 codieren eine Acyltransferase-Domäne;

Nucleotidpositionen 2146-2697 der SEQ ID NO: 6, Aminosäurepositionen 4447-4630 codieren eine Dehydratase-Domäne;

25 Nucleotidpositionen 3604-4155 der SEQ ID NO: 6, Aminosäurepositionen 4933-5116 codieren eine Ketoreduktase-Domäne;

Nucleotidpositionen 4420-4677 der SEQ ID NO: 6, Aminosäurepositionen 5205-5290 codieren eine  $\beta$ -Ketoacyl:Acyl Carrier Protein-Domäne;

Nucleotidpositionen 4864-5538 der SEQ ID NO: 6, Aminosäurepositionen 5353-5577 codieren eine Thioesterase-Domäne.

SEQ ID NOS: 52 und 53, ORF23:

Nucleotidposition 344 bis 1333 der SEQ ID NO: 51, 329 Aminosäuren;  
das ableitbare Genprodukt ist eine dNDP-Glucose-4,6-Dehydratase.

5 SEQ ID NOS: 54 und 55, ORF24:

Nucleotidposition 1330 bis 2247 der SEQ ID NO: 51, 305 Aminosäuren;  
das ableitbare Genprodukt ist eine dNDP-4-keto-6-Deoxyglucose-3,5-Epimerase.

10 Die an der Cyclisierung des 5, 6, 5- Tricyclus beteiligten Produkte des ORF 5 (SEQ  
ID NO: 16) und des ORF 8 (SEQ ID NO: 22) sind aufgrund der ungewöhnlichen  
Cyclisierungsreaktionen von besonderem Interesse. Daher beinhaltet die vorliegende  
Erfindung insbesondere auch homologe Nucleinsäuren oder homologe Genprodukte.  
Vorzugsweise zeigen diese homologen Genprodukte mindestens eine 50 %ige,  
bevorzugt eine 60 %ige und besonders bevorzugt eine 70 %ige Identität auf Amino-  
15 säureebene.

Weiterhin sind Antikörper Gegenstand der Erfindung, die spezifisch an die  
vorstehend genannten Polypeptide binden. Die Herstellung solcher Antikörper erfolgt  
auf die übliche Weise. Diese Antikörper können genutzt werden, um Expres-  
20 sionsklone z.B. einer Genbank zu identifizieren, die die erfindungsgemäßen Nuclein-  
säuren tragen.

Gegenstand der vorliegenden Erfindung sind auch Verfahren zum Herstellen der  
erfindungsgemäßen Nucleinsäuren. Die erfindungsgemäßen Nucleinsäuren können  
25 auf die übliche Weise hergestellt werden. Beispielsweise können die Nucleinsäure-  
moleküle vollständig chemisch synthetisiert werden. Man kann auch kurze Stücke  
der erfindungsgemäßen Nucleinsäuren chemisch synthetisieren und solche

Oligonucleotide radioaktiv oder mit einem Fluoreszenzfarbstoff markieren. Die  
markierten Oligonucleotide können auch verwendet werden, um Genbanken von  
30 Organismen zu durchsuchen. Klone, an die die markierten Oligonucleotide hybridi-  
sieren, werden zur Isolierung der betreffenden DNA ausgewählt. Nach der Charak-

terisierung der isolierten DNA erhält man auf einfache Weise die erfindungsgemäßen Nucleinsäuren. Die erfindungsgemäßen Nucleinsäuren können auch mittels PCR-Verfahren unter Verwendung chemisch synthetischer Oligonucleotide hergestellt werden.

5

Gegenstand der vorliegenden Erfindung sind weiterhin Verfahren zum Herstellen der erfindungsgemäßen Polypeptide. Zur Herstellung der Polypeptide, die von den erfindungsgemäßen Nucleinsäuren codiert werden, können Wirtszellen, die zumindest eine der erfindungsgemäßen Nucleinsäuren enthalten, unter geeigneten Bedingungen kultiviert werden. Die gewünschten Polypeptide können danach auf übliche Weise aus den Zellen oder dem Kulturmedium isoliert werden. Die Polypeptide können auch in *in vitro*-Systemen hergestellt werden.

10

15

Das isolierte und charakterisierte Gencluster und benachbarte oder assoziierte DNA-Regionen stellen ein Target zur Steigerung der Spinosyn-Biosynthese durch genetische Manipulation, Über- oder Unterexpression von direkt oder indirekt an der Biosynthese involvierten Genen oder regulatorischen Sequenzen dar. Diese Manipulationen können sowohl in natürlichen Spinosyn-produzierenden Organismen als auch in gentechnisch hergestellten Spinosyn-produzierenden Organismen durchgeführt werden. Beispielsweise können ausgewählte ORF's unter die Kontrolle üblicher starker Promotoren wie dem *mel*-Promotor des Plasmides pIJ702 (John Innes Foundation, Norwich, UK, 1985) gestellt werden.

20

25

Durch die Klonierung und Identifizierung der Spinosyn-Biosynthesegene schafft die vorliegende Erfindung die genetische Basis, mittels molekulargenetischer Verfahren neue Spinosyn-Vorstufen und -Derivate herzustellen.

---

Der Ausdruck "Spinosyn-Vorstufen", wie er hierin verwendet wird, bezieht sich auf alle nachweisbaren und alle postulierbaren Biosynthesevorstufen von Spinosyn.

30



Der Ausdruck "Spinosyn-Derivate", wie er hierin verwendet wird, bezieht sich auf Strukturderivate aller bisher bekannten Spinosyne.

5 Gegenstand der vorliegenden Erfindung ist somit auch ein Verfahren zum Herstellen von Spinosyn-Vorstufen und -Derivaten.

Die erfindungsgemäßen Nucleinsäuren können beispielsweise eingesetzt werden, um durch kombinatorische Biosynthese neue Spinosyn-Derivate mit Veränderungen des Spinosyn-Aglycons herzustellen. Dies kann z.B. dadurch erreicht werden, dass die von ORF 19 codierte, eine Acetat-Einheit einbauende Acyltransferase-Domäne ausgetauscht wird gegen eine Acyltransferase-Domäne, die eine Propionat-Einheit einbaut. In gleicher Weise kann die, eine Acetat-Einheit einbauende Acyltransferase-Domäne des ORF 18 gegen eine Acyltransferase-Domäne ausgetauscht werden, die eine Propionat-Einheit einbaut. Ferner ist es möglich beide oder jeweils eine Ketoreduktase-Domäne, die von beiden genannten ORF's codiert werden zu inaktivieren, durch eine inaktive Ketoreduktase-Domäne zu ersetzen oder zu deletieren, wodurch eine Hydroxygruppe an der entsprechenden Position im Makrocyclus biosynthetisch hergestellt werden kann. Alle Acyltransferase-, Ketoreduktase-, Dehydratase-, Enoylreduktase-,  $\beta$ -Ketoacyl:Acyl Carrier Protein und Thioesterase-Domänen können einzeln oder in beliebiger Kombination durch entsprechende Polyketidsynthase-Domänen mit anderer Substrat- oder Reaktionsspezifität ersetzt werden, in beliebiger Kombination miteinander fusioniert, einzeln oder in beliebiger Kombination mutagenisiert, deletiert oder dupliziert werden. Ferner können Modul-codierende Sequenzen ausgetauscht werden. So ist es denkbar die Modul 2-codierende DNA-Sequenz (Abb. 6) gegen die Modul 1- oder Modul 3, 4, 5, 6, 7, 8- oder Modul 9-codierende DNA-Sequenz (Abb. 6) zu ersetzen und funktionell zu exprimieren. Es ist auch denkbar die Modul 2-codierende DNA-Sequenz oder jede andere Modul-codierende DNA-Sequenz des Spinosyn-Polyketidsynthase-Genclusters gegen eine andere Modul-codierende DNA-Sequenz des Spinosyn-Polyketidsynthase-Genclusters, die eine andere biosynthetische Verlängerungseinheit einbaut, auszutauschen. Darüber hinaus kann jede andere Modul-codierende DNA-

Sequenz des Spinosyn-Polyketidsynthase-Genclusters gegen eine andere Modul-codierende DNA-Sequenz einer anderen Polyketidsynthase-Nukleinsäuresequenz aus *S. spinosa* oder einem anderen Organismus als *S. spinosa*, wie z.B. *Saccharopolyspora erythraea*, ausgetauscht werden.

5

Gegenstand der Erfindung sind somit auch alle Modul- oder Domänen-codierenden Nucleinsäuren, die natürlicher oder gentechnisch erzeugter Bestandteil der Spinosyn-Polyketidsynthase sind.

10

15

20

Der Ausdruck "Modul", sowie er hierin verwendet wird, bedeutet, dass eine Anordnung von drei enzymkatalytisch aktiven Domänen vorliegt, die zu eine Verlängerung der wachsenden Polyketidkette um eine biosynthetische Verlängerungseinheit führen. Bei diesen Domänen handelt es sich um eine  $\beta$ -Ketoacyl:Acyl Carrier Protein Synthase-Domäne, eine Acyltransferase-Domäne und eine  $\beta$ -Ketoacyl:Acyl-Carrier Protein-Domäne. Ein Modul kann auch eine Ketoreduktase-, eine Dehydratase-, eine Enoylreduktase- und eine Thioesterase-Domäne tragen. Ein sog. Ladungsmodul, das am Beginn der Biosynthese steht kann von den genannten Domänen lediglich eine Acyltransferase-Domäne und eine  $\beta$ -Ketoacyl:Acyl-Carrier Protein-Domäne tragen, sowie eine enzymatisch inaktive  $\beta$ -Ketoacyl:Acyl Carrier Protein Synthase-Domäne. Eine Polyketidsynthase-Domäne umfasst jeweils eine dieser genannten enzymatischen Aktivitäten.

25

30

Die erfindungsgemäßen Nucleinsäuren können weiterhin genutzt werden, um im Zuge einer kombinatorischen Biosynthese durch Neuordnung und Expression von Spinosyn-Polyketidsynthase-Nucleinsäuresequenzen oder durch Kombination und Expression zusammen mit Polyketidsynthase-Nucleinsäuresequenzen einer anderen Polyketidsynthase codierenden Nucleinsäuresequenz aus *S. spinosa* oder einem anderen Organismus, wie z.B. *Saccharopolyspora erythraea*, Bibliotheken von rekombinanten Polyketidsynthase-Nucleinsäuresequenzen, rekombinanten Polyketidsynthase-Proteinen oder rekombinant erzeugten Polyketiden herzustellen. Diese Polyketide können durch die Verwendung der erfindungsgemäßen Nucleinsäuren

oder die Verwendung anderer Nucleinsäuren, deren ableitbaren Produkte an der Biosynthese anderer Zucker und Ankopplung ans Aglycon beteiligt sind, glycosyliert werden. Es ist bekannt, dass die Glycosylierung des Aglycons eine entscheidende Rolle bei der biologischen Aktivität am Wirkort spielt. Diese Veränderungen können sowohl in natürlichen als auch in gentechnisch hergestellten Spinosyn-produzierenden Organismen, insbesondere Bakterien, erfolgen.

Die erfindungsgemäßen Nucleinsäuren, Vektoren und regulatorischen oder genetisch mobilen Regionen können außerdem zum Auffinden von Genen verwendet werden, die für Polypeptide codieren, welche funktionell ähnliche Polyketidsynthasen oder funktionell ähnliche Produkte, die an einer Zuckerbiosynthese beteiligt sind, codieren.

Da die erfindungsgemäßen Nucleinsäuren einen umfangreichen Teil des Genoms von *S. spinosa* ausmacht, können die erfindungsgemäßen Nucleinsäuren als Marker bei der Sequenzierung des Genoms von *S. spinosa* eingesetzt werden, wodurch die Anordnung von Teilsequenzen eines Genomsequenzierungsprojektes erheblich erleichtert wird.

Somit liefern die erfindungsgemäßen Nucleinsäuren Daten, die im Rahmen eines Genomsequenzierungsprojektes und eines sich darauf aufbauenden Metabolic Engineering zur Steigerung der Spinosynproduktion eingesetzt werden können.

#### Erläuterungen zu den Abbildungen:

Abbildung 1: Modell für die Biosynthese der Spinosyn-Zucker D-Forosamin und 2, 3, 4-Tri-O-Methyl-L-Rhamnose.

Abbildung 2: Lage, der an der Spinosyn-Biosynthese direkt oder indirekt beteiligten DNA-Regionen 1 (SEQ ID NO: 4) und DNA-Region 2 (SEQ ID NOS: 5 und 6). Die schwarzen Balken im unteren Teil der Abbildung geben schematisch die Positionen

der Cosmid-DNA Inserts zueinander und in Bezug zu den DNA-Regionen 1 und 2 an. Die dargestellten Cosmid-Inserts wurden zur Sequenzierung der SEQ ID NOS: 1 bis 3 herangezogen.

5 Abbildung 3: Schematische Darstellung der Lage der Insert-DNA (schwarze Balken im unteren Teil der Abbildung) der benannten Cosmide, die zur Ankopplung eines Forosamin-Restes oder eines Trimethyl-Rhamnose-Restes durch Biotransformation des Spinosyn-Aglycons und Spinosyn-Pseudoaglycons herangezogen worden sind.

10 Abbildung 4: Schematische Darstellung der offenen Leserahmen (ORF's) der DNA-Region 3, die der SEQ ID NO: 51 entspricht, auf Cosmid 16-2-2.

15 Abbildung 5: Schematische Darstellung offener Leserahmen (ORF's) der DNA-Regionen 1 und 2. Die ORF's sind nummeriert von 1 bis 22 entsprechend SEQ ID NOS: 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47 und 49.

20 Abbildung 6: Schematische Darstellung offener Leserahmen (ORF's) der DNA-Region 2 (SEQ ID NOS: 5 und 6) und ableitbarer Module und Domänen. SM, Startmodul; M1 bis M10, Modul 1 bis Modul 10; KS,  $\beta$ -Ketoacyl:Acyl Carrier Protein Synthase; AT, Acyltransferase; ACP,  $\beta$ -Ketoacyl:Acyl Carrier Protein; KR, Ketoreduktase; DH, Dehydratase; ER, Enoylreduktase.

25 Abbildung 7: Schematische Darstellung der Lage von BAC-Shuttleklon Insert-DNA als schwarze Balken im unteren Teil der Abbildung. Die Größe der Insert-DNA beträgt mindestens 100 kb. Durchgezogene Balken: DNA-Sequenz ist identisch mit ~~Teilen bzw. der Gesamtheit der DNA-Regionen 1 und 2.~~ Gestrichelte Balken: DNA-Sequenz liegt außerhalb des sequenzierten Bereichs.

## Beispiele

### Bakterienstämme und Plasmide

5 *Escherichia coli* XL1-Blue MRF' und die Cosmidvektoren SuperCos1 (Stratagene, Europe) und pOJ446 (Biermann et al., 1992) wurden verwendet zur Erstellung von Genbanken von *S. spinosa* ATCC49460 (American Type Culture Collection, U.S.A., EP-A 0 375 316). *E. coli* JM110 (Stratagene, Europe) wurde verwendet zur Propagierung von Plasmiden, die durch Transformation nach *Streptomyces* übertragen wurden. *Streptomyces albus* J1074 (Chater and Wilde, 1980; John Innes Institut in  
10 Norwich, UK) wurde zur heterologen Expression von und zur Biotransformation mit Spinosyn-Biosynthesegenen eingesetzt.

Plasmid pBeloBAC11 (Kim et al., 1996) und pOJ446 (Biermann et al., 1992) wurden  
15 verwendet zur Herstellung eines *E. coli* – *Streptomyces* BAC-Shuttlevektors.

### Molekularbiologische Methoden

20 Molekularbiologische Methoden wie DNA-Restriktion, Agarose-Gelelektrophorese von DNA, Ligation von Restriktionsfragmenten, Kultivierung und Transformation von *E. coli* wurden durchgeführt wie beschrieben in Sambrook et. al (1989). Plasmide wurden mit Qiagen Plasmid Kit (Qiagen, Hilden, Deutschland) isoliert. Die verwendeten Enzyme stammten von Roche Diagnostics GmbH (Mannheim, Deutschland).

25 Anzucht Bedingungen and molekulargenetische Methoden mit *S. spinosa* und Streptomyceten sind beschrieben in (Hopwood et al., 1985). Alle Anzuchten in Flüssigkultur von *S. spinosa* oder Streptomyceten erfolgten aerob in Erlenmeyerkolben bei 28°C.

---

Die DNA-DNA-Hybridisierungen erfolgten unter Verwendung des DIG-High Prime DNA Labeling and Detection Kit nach Angaben des Herstellers (Roche Diagnostics GmbH, Mannheim, Deutschland).

5 Wachstumsmedien:

LB Sambrook et. al., 1989

10 TS Difco Bestell-Nummer 0 370-17-3 (Difco Detroit, MI, USA)

R5A Illing et al., 1985

Herstellung einer Cosmid Genbank von *S. spinosa*

15 Um eine Genbank von *S. spinosa* zu erhalten, wurde chromosomale DNA von *S. spinosa* ATCC49460 mit *Mbo*I partiell geschnitten und durch Zentrifugation im Glucosedichtegradienten aufgetrennt. Die Cosmid-DNA (SuperCos1, Stratagene Europe) wurde nach Angaben des Herstellers vorbereitet, mit den *S. spinosa* DNA-Fragmenten zwischen 35 und 45 kb ligiert und mit Hilfe des Gigapack-Verpackungssystem (Stratagene Europe) in Phagenpartikel verpackt. Die Transfektion erfolgte in *E. coli* XL-1 blue MRF'. Diese Methode wurde ebenfalls dazu verwendet  
20 eine zweite *S. spinosa* Genbank anzulegen unter Verwendung des *E. coli*-*Streptomyces* Shuttle-Cosmids pOJ446.

25 Sequenzierung des Spinosynbiosynthese-Genclusters und eines DNA-Fragmentes das außerhalb dieses Clusters liegt, dessen Produkte aber an der Biosynthese von Spinosyn beteiligt sind

30 Die Insert-DNA der SuperCos1 Cosmide 16-1-8, 16-59-1, und 16-59-8 wurden sequenziert. Eine ca. 4 kb große Lücke zwischen den Cosmiden 16-59-1 und 16-1-8

wurde durch primer walking Sequenzierung eines entsprechenden Teilbereiches von Cosmid 16-59-6 geschlossen.

Eine ca. 2,3 kb große DNA-Sequenz auf dem SuperCos1 Cosmid 16-2-2 wurde sequenziert.

Identifizierung und Charakterisierung von chromosomalen DNA-Fragmenten einer BAC-Shuttlevektor-Genbank aus *S. spinosa*, die Spinosyn-Biosynthese-Gensequenzen tragen

Zur Herstellung des BAC-Shuttlevektors, der nicht nur in *E. coli* sondern auch in Actinomyceten wie *Streptomyces* übertragen und vermehrt werden kann, wurde der Vektor pBeloBAC11 mit *Xho*I linearisiert, und durch die Anwendung von Klenow Polymerase wurden glatte DNA-Enden hergestellt. Ein ca. 6 kb großes *Dra*I – *Eco*RV DNA-Fragment des Cosmidvektors pOJ446, das den Replikationsursprung des Plasmides SCP2\*, das Apramycinresistenzgen sowie den *oriT* zum konjugativen Transfer trägt, wurde mit dem linearisierten BAC-Vektor ligiert. Der resultierende Vektor erhielt die Bezeichnung pEBZ333.

Ausgehend von partiell mit *Mbo*I geschnittener genomischer DNA des Stammes *S. spinosa* ATCC49460 und dem mit *Bam*HI geschnittenen Vektor pEBZ333 wurde eine BAC-Genbank erstellt.

Analyse und Annotation offener Leserahmen direkt oder indirekt an der Spinosyn-Biosynthese beteiligter DNA-Sequenzen

~~Ausgehend von der Sequenz gemäß SEQ ID NOS: 1 bis 3 wurden offene Leserahmen (ORF's) identifiziert, die direkt oder indirekt an der Biosynthese von Spinosyn beteiligt sind. Diese ORF's wurden in zwei DNA-Regionen unterteilt, die als DNA-Region 1 und DNA-Region 2 (Abb. 2 und 5) bezeichnet werden und die~~

Sequenzen gemäß SEQ ID NO: 4 bzw. 5 und 6 tragen. Die DNA-Region 1 trägt offene Leserahmen, deren Produkte an der Modifizierung und Tricyclusbildung des Spinosyn-Aglycons beteiligt sind, während die DNA-Region 2 (Abb. 2, 5 und 6) offene Leserahmen umfasst, deren Produkte die Spinosyn-Polyketidsynthase codieren. Die beiden jeweils ersten Nucleotide dieser DNA-Regionen liegen unmittelbar nebeneinander (Abb. 2, 3 und 5).

Eine weitere DNA-Region 3 (SEQ ID NO: 51) liegt außerhalb dieses Clusters von DNA-Sequenzen und trägt offene Leserahmen, deren Produkte ebenfalls an der Biosynthese des Spinosyn-Zuckers Trimethyl-Rhamnose beteiligt sind.

#### Herstellung des Spinosyn-Aglycons und 17-Pseudoaglycons aus Tracer®

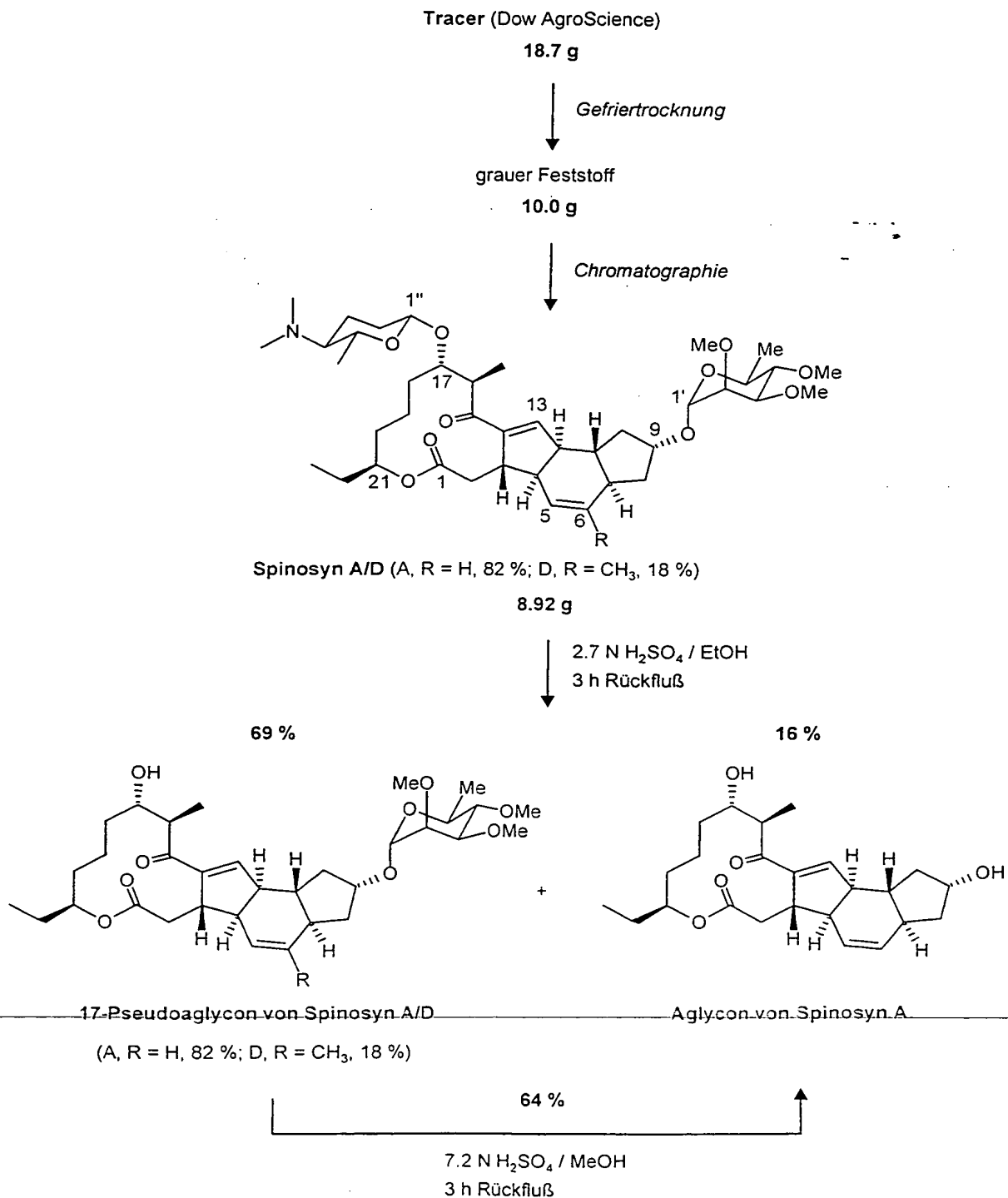
Ausgehend von 18.7 g des kommerziell erhältlichen Produktes Tracer® wurden nach Gefriertrocknung und Säulenchromatographie an Kieselgel 8.92 g Spinosyn A und D in einem Verhältnis von 82:18 gewonnen.

Die Hydrolyse des Aminozuckers Forosamin gelang mit 2.7 N Schwefelsäure in Ethanol unter Rückfluß. Dabei fiel der Großteil des entstehenden 17-Pseudoaglycons von Spinosyn A/D aus. Im Filtrat wurden neben weiterem 17-Pseudoaglycon in Abhängigkeit von der Reaktionsdauer geringe bis mittlere Mengen des Spinosyn-Aglycons gefunden.

Eine vollständige Hydrolyse zum Aglycon gelang unter etwas drastischeren Bedingungen (7.2 N Schwefelsäure in Methanol unter Rückfluß). Die Aglycon-Fraktion enthielt ausschließlich Aglycon von Spinosyn A. Dies stimmt sehr gut mit der Literatur überein (Creemer et al., 1998), die unter entsprechenden Reaktionsbedingungen eine vollständige Zersetzung des Pseudoaglycons von Spinosyn D beschreibt. Als Ursache nehmen die Autoren eine leichtere Protonierung der 5,6-Doppelbindung bei Spinosyn D unter Bildung eines tertiären Carbokations und anschließende Umlagerungen an.



Es konnten somit ausgehend von 18.7 g käuflichem Tracer® 3.0 g Aglycon von Spinosyn A hergestellt werden.



## Gewinnung von Spinosyn A/D aus Tracer®

Die Gefriertrocknung von 18,7 g Tracer® lieferten 10,0 g grauen Feststoff. Nach Säulenchromatographie dieses Feststoffes an 800 cm<sup>3</sup> Kieselgel (Eluent: Dichlormethan/Methanol 95:5) erhielt man 8,92 g reines Spinosyn A/D (82 % A, 18 % D). – DC:  $R_f$  (SiO<sub>2</sub>, Dichlormethan/Methanol 9:1) = 0,46. – <sup>1</sup>H-NMR: CDCl<sub>3</sub>,  $\delta$  = 6,77 (s, 13-H); 5,88 (d, 5-H von Spinosyn A); 5,80 (m, 6-H von Spinosyn A); 5,49 (m, 5-H von Spinosyn D); 4,87 (d, 1'-H); 4,67 (m, 21-H); 4,43 (d, 1''-H); 4,31 (m, 9-H) u. a. – LC/MS: Elektrospray Positiv; Peak bei RT 44,0 min:  $m/z$  = 733 (100 %) [M+H]<sup>+</sup> (Spinosyn A); Peak bei 44,7 min:  $m/z$  = 747 (100 %) [M+H]<sup>+</sup> (Spinosyn D).

Darstellung des 17-Pseudoaglycons von Spinosyn A/D:

8,65 g (11,81 mmol) Spinosyn A/D wurden in 61 ml Ethanol gelöst und mit 104 ml Wasser und 208 ml 4 N H<sub>2</sub>SO<sub>4</sub> versetzt. Nach 3 h Erwärmen unter Rückfluß wurde der ausgefallene Feststoff (A) abfiltriert und getrennt vom Filtrat (B) aufgearbeitet. Der Feststoff (A) wurde mit 1 N H<sub>2</sub>SO<sub>4</sub> gewaschen, in 140 ml Dichlormethan aufgenommen, nacheinander mit gesätt. NaHCO<sub>3</sub>-Lösung und gesätt. NaCl-Lösung gewaschen, über Na<sub>2</sub>SO<sub>4</sub> getrocknet und im Vakuum eingeeengt. Umkristallisation aus Ethanol lieferten 3,03 g 17-Pseudoaglycon von Spinosyn A/D und Mutterlauge (C). Das Filtrat (B) wurde mehrmals mit Dichlormethan extrahiert. Die Extrakte wurden nacheinander mit gesätt. NaHCO<sub>3</sub>-Lösung und gesätt. NaCl-Lösung gewaschen, über Na<sub>2</sub>SO<sub>4</sub> getrocknet und im Vakuum eingeeengt. Der Rückstand wurde vereint mit der im Vakuum eingeeengten Mutterlauge (C) und durch Säulenchromatographie an 650 cm<sup>3</sup> Kieselgel (Eluent: Cyclohexan/Essigsäureethylester 1:1, dann 100 % Essigsäureethylester) aufgetrennt. Man erhielt neben weiteren 1,76 g 17-Pseudoaglycon von Spinosyn A/D 0,78 g (16 %) Aglycon von Spinosyn A. Die Gesamtausbeute von 17-Pseudoaglycon von Spinosyn A/D betrug 4,79 g (69 %). – a) 17-Pseudoaglycon von Spinosyn A/D (82 % A, 18 % D): DC:  $R_f$  (SiO<sub>2</sub>, Essigsäureethylester) = 0,48. – <sup>1</sup>H-NMR: CDCl<sub>3</sub>,  $\delta$  = 6,78 (s, 13-H); 5,88 (d, 5-H von Spinosyn A); 5,80 (m, 6-H von Spinosyn A); 5,49 (m, 5-H von Spinosyn D); 4,86 (d,

1'-H); 4,70 (m, 21-H); 4,32 (m, 9-H) u. a. - LC/MS: Elektrospray Positiv; Peak bei RT 40,7 min:  $m/z = 609$  (100 %)  $[M+NH_4]^+$ ,  $m/z = 641$  (10 %)  $[M+NH_4+CH_3OH]^+$  (Pseudoaglycon von Spinosyn A); Peak bei RT 41,4 min:  $m/z = 623$  (100 %)  $[M+NH_4]^+$ ,  $m/z = 655$  (8 %)  $[M+NH_4+CH_3OH]^+$  (Pseudoaglycon von Spinosyn D). –  
 5 b) Aglycon von Spinosyn A: DC:  $R_f$  (SiO<sub>2</sub>, Essigsäureethylester) = 0,29. – <sup>1</sup>H-NMR: CDCl<sub>3</sub>,  $\delta = 6,80$  (s, 13-H); 5,89 (d, 5-H); 5,80 (m, 6-H); 4,70 (m, 21-H); 4,44 (m, 9-H) u. a. - LC/MS: Elektrospray Positiv; Peak bei RT 36,8 min:  $m/z = 420$  (100 %)  $[M+NH_4]^+$ ,  $m/z = 452$  (10 %)  $[M+NH_4+CH_3OH]^+$ .

#### 10 Darstellung des Aglycon von Spinosyn A/D

4,30 g (7,29 mmol) Pseudoaglycon von Spinosyn A/D wurden in 190 ml Methanol gelöst und mit 285 ml 7,2 N H<sub>2</sub>SO<sub>4</sub> versetzt. Nach 3 h Erwärmen unter Rückfluß wurde die abgekühlte Reaktionsmischung vorsichtig in 1700 ml gesättigte NaHCO<sub>3</sub>-  
 15 Lösung gegeben. Man extrahierte mit Diethylether, wusch die Extrakte nacheinander mit gesätt. NaHCO<sub>3</sub>-Lösung und gesätt. NaCl-Lösung, trocknete über Na<sub>2</sub>SO<sub>4</sub> und engte im Vakuum ein. Nach Säulenchromatographie dieses Feststoffes an 650 cm<sup>3</sup> Kieselgel (Eluent: Cyclohexan/Essigsäureethylester 1:2, dann 100 % Essigsäureethylester) erhielt man 1,88 g (64 %) Aglycon von Spinosyn A. – DC:  $R_f$  (SiO<sub>2</sub>, Essigsäureethylester) = 0,29. – <sup>1</sup>H-NMR: CDCl<sub>3</sub>,  $\delta = 6,80$  (s, 13-H); 5,89 (d, 5-H); 5,80 (m, 6-H); 4,70 (m, 21-H); 4,44 (m, 9-H) u. a. – LC/MS: Elektrospray Positiv; Peak bei RT 36,6 min:  $m/z = 420$  (100 %)  $[M+NH_4]^+$ ,  $m/z = 452$  (14 %)  $[M+NH_4+CH_3OH]^+$  (Aglycon von Spinosyn A).  
 20

25 Forosaminylierung des Spinosyn-Aglycons und Anknüpfung eines Trimethyl-Rhamnosezuckers an das Spinosyn-Aglycon durch Biotransformation mit einem rekombinanten *Streptomyces* Stamm, der Spinosyn-Zuckerbiosynthesegene heterolog exprimiert

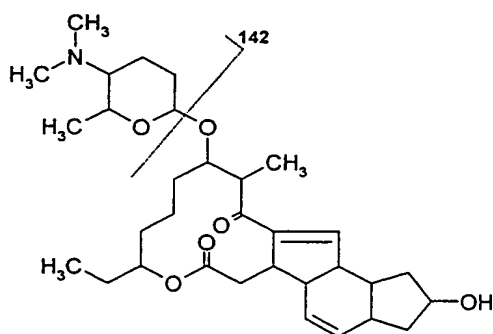
30 20 ml R5A Medium (Illing et al., 1989) mit 5 µg Apramycin / ml wurde mit Mycel des rekombinanten Stammes *S. albus* (165-1) oder *S. albus* (165-8) angeimpft und

24 h aerob bei 28°C bebrütet. Diese Kultur wurde mit 50 µg/ml des hergestellten Spinosyn-Aglycons (100 µl einer 1 %igen Stammlösung in Methanol; Herstellung siehe Kapitel "Herstellung des Spinosyn-Aglycons und 17-Pseudoaglycons aus Tracer®", „Gewinnung von Spinosyn A/D aus Tracer®“ und „Darstellung des Aglycon von Spinosyn A/D“) versetzt und ca. 120 h bei 28°C aerob inkubiert. Als Kontrolle wurde in gleicher Weise *S. albus* (pEBZ340; pOJ446-Vektor mit einem ca. 1,8 kb großen Spinosyn-PKS tragenden DNA-Fragment aus Cosmid 16-1-8) kultiviert und mit Spinosyn-Aglycon versetzt. Die Kulturen wurden nach Inkubation zur Abtrennung von Zellmycel zentrifugiert und der Überstand (20 ml) wurde mit 25 ml Methanol versetzt.

Je 35 ml des mit Methanol versetzten Kulturüberstandes wurden lyophilisiert, mit 15 ml Wasser aufgenommen und zweimal mit je 10 ml Essigsäureethylester extrahiert. Die vereinten organischen Phasen wurden zur Trockene eingengt und mit 350 µl Methanol aufgenommen. Ein Aliquot dieser Extrakte wurde mittels LC/MS mit Elektrospray Positiv-Ionisation untersucht.

Der Kulturüberstand der Anzucht von *S. albus* (165-1) enthielt eine Verbindung mit dem Molekulargewicht eines forosaminylierten Aglycons von Spinosyn A sowie Spinosyn A.

Peak 1: RT = 41,0 min: m/z = 544 (100 %) [M+H]<sup>+</sup>, m/z = 576 (20 %) [M+H+CH<sub>3</sub>OH]<sup>+</sup> (Forosaminyliertes Aglycon von Spinosyn A); LC/MS/MS: m/z = 142 (30 %) (Forosamin-Fragment).



Forosaminyliertes Aglycon von Spinosyn A  
 Molekulargewicht = 543  
 Summenformel =  $C_{32}H_{49}NO_6$

Peak 2: RT = 44,2 min:  $m/z = 733$  (100 %)  $[M+H]^+$  (Spinosyn A); LC/MS/MS:  $m/z = 142$  (21 %) (Forosamin-Fragment).

5

Der Kulturüberstand von *S. albus* (165-8) enthielt eine Verbindung mit dem Molekulargewicht eines Forosaminylierten Aglycons von Spinosyn A.

Peak 1: RT = 40,9 min:  $m/z = 544$  (100 %)  $[M+H]^+$ ,  $m/z = 576$  (20 %)  $[M+H+CH_3OH]^+$  (Forosaminyliertes Aglycon von Spinosyn A); LC/MS/MS:  $m/z = 142$  (39 %) (Forosamin-Fragment).

10

Der Kulturüberstand von *S. albus* (pEBZ340) enthielt keine Verbindungen mit MW 544 und kein Spinosyn A.

Forosaminylierung des Spinosyn-17-Pseudoaglycons durch Biotransformation mit einem rekombinanten *Streptomyces* Stamm, der Spinosyn Zuckerbiosynthesegene heterolog exprimiert

15

20 ml R5A Medium (Illing et al., 1989) mit 5  $\mu$ g Apramycin / ml wurde mit Mycel des rekombinanten Stammes *S. albus* (165-1) oder *S. albus* (165-8) angeimpft und 24 h aerob bei 28°C bebrütet. Diese Kultur wurde mit 50  $\mu$ g/ml des hergestellten 17-Pseudoaglycons von Spinosyn (100  $\mu$ l einer 1 %igen Stammlösung in Methanol; Herstellung siehe Kapitel "Herstellung des Spinosyn-Aglycons und 17-Pseudo-

20

aglycons aus Tracer<sup>®</sup>, „Gewinnung von Spinosyn A/D aus Tracer<sup>®</sup>“ und „Darstellung des 17-Pseudoaglycon von Spinosyn A/D““) versetzt und ca. 120 h bei 28°C aerob inkubiert. Die Kulturen wurden nach Inkubation zur Abtrennung von Zellmycel zentrifugiert und der Überstand (20 ml) wurde mit 25 ml Methanol versetzt.

Je 35 ml des mit Methanol versetzten Kulturüberstandes wurden lyophilisiert, mit 15 ml Wasser aufgenommen und zweimal mit je 10 ml Essigsäureethylester extrahiert. Die vereinten organischen Phasen wurden zur Trockene eingedunstet und mit 350 µl Methanol aufgenommen. Ein Aliquot dieser Extrakte wurde mittels LC/MS mit Elektrospray Positiv-Ionisation untersucht.

Der Kulturüberstand von *S. albus* (165-1) enthielt Spuren von Spinosyn A und D.  
Peak 1: RT = 44,2 min: m/z = 733 (100 %) [M+H]<sup>+</sup> (Spinosyn A); LC/MS/MS: m/z = 142 (8 %) (Forosamin-Fragment).  
Peak 2: RT = 44,7 min: m/z = 747 (100 %) [M+H]<sup>+</sup> (Spinosyn D); LC/MS/MS: m/z = 142 (37 %) (Forosamin-Fragment).

Der Kulturüberstand von *S. albus* (165-8) enthielt Spuren von Spinosyn A und D.  
Peak 1: RT = 44,1 min: m/z = 733 (100 %) [M+H]<sup>+</sup> (Spinosyn A).  
Peak 2: RT = 44,7 min: m/z = 747 (100 %) [M+H]<sup>+</sup> (Spinosyn D).

#### Hinterlegung von Mikroorganismen

Folgende Mikroorganismen und Plasmide sind bei der Deutschen Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ), Mascheroder Weg 1b, D- 38124 Braunschweig, in Übereinstimmung mit den Bestimmungen des Budapester Vertrages hinterlegt worden.

## Mikroorganismus und Plasmid

## Hinterlegungsnummer

	<i>E. coli</i> XL1-Blue MRF' mit Cosmid 16-1-8	DSM 12961
	<i>E. coli</i> XL1-Blue MRF' mit Cosmid 16-2-2	DSM 12962
	<i>E. coli</i> XL1-Blue MRF' mit Cosmid 16-59-1	DSM 12963
5	<i>E. coli</i> XL1-Blue MRF' mit Cosmid 16-59-6	DSM 12964
	<i>E. coli</i> XL1-Blue MRF' mit Cosmid 16-59-8	DSM 12965
	<i>E. coli</i> XL1-Blue MRF' mit Cosmid 165-1	DSM 13005
	<i>E. coli</i> XL1-Blue MRF' mit Cosmid 165-8	DSM 13007
	<i>E. coli</i> DH10B mit dem BAC Shuttle-Klon P8 / G11	DSM 13012
10	<i>E. coli</i> DH10B mit dem BAC Shuttle-Klon P11 / B10	DSM 13011
	<i>E. coli</i> DH10B mit dem BAC Shuttle-Klon P11 / G6	DSM 13010

Literatur

- 15 Biermann M., Logan R., O'Brien K., Seno E.T., Nagaraja R., Schoner B.E. (1992) Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp. Gene 116: 43-49

- 20 Chater K., Wilde L. (1980) *Streptomyces albus* G Mutants defective in the *SalGI* restriction-modification system. J. Gen. Microbiol. 116: 323-334.

Devereux J., Haeberli P., Smithies (1984) A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Research 12: 387-395.

- 25 Mertz F., Yao R.C. (1990) *Saccharopolyspora spinosa* sp. nov. isolated from soil collected in a sugar mill rum still. Int. J. Syst. Bacteriol. 40: 34-39.

- 
- 30 Creemer L.C., Kirst H.A., Paschal J.W. (1998) Conversion of spinosyn A and spinosyn D to their respective 9- and 17-pseudoaglycones and their aglycones. J. Antibiotics 51: 795-800.

Fishman S.E., Hersberger C.L. (1983) Amplified DNA in *Streptomyces fradiae*. J. Bacteriol. 155: 459-466.

5 Hopwood D.A., Bibb M.J., Chater K.F., Kieser T., Bruton C.J., Kieser H.M., Lydiate D.J., Smith C.P., Ward J.M., Schrempf H. (1985) Genetic manipulation of *streptomyces* a laboratory manual, The John Innes Foundation, Norwich, 1985.

10 Illing G.T., Normansell I.D., Peberdy J.F. (1989) Protoplast isolation and regeneration in *Streptomyces clavuligerus*. J. Gen. Microbiol. 135: 2289-2297.

Kim U.J., Birren B.W., Slepak T., Mancino V., Boysen C., Kang H.L., Simon M.I., Shizuya H. (1996), Construction and characterization of a human bacterial artificial chromosome library. Genomics 34: 213-218.

15 Kirst H.A., Michel K.H., Martin J.W., Creemer L.C., Chao E.H., Yao R.C., Nakatsukasa W.M., Boeck L.D., Occolowitch J.L., Paschal J.W., Deeter J.B., Jones N.D., Thompson G.D. (1991) Tetrahedron Lett. 32: 4839-4842

20 Nakatsukasa W.M., Mabe J.A., Michel K.H., Martin J.W., Paschal J.W., Elzey T.K. (1990) Abstracts of 2<sup>nd</sup> Int. Conf. on Biotechnology of Microbial Prod.: P-21.

Sambrook J., Fritsch E.F., Maniatis T. (1989) Molecular cloning - A laboratory manual. Second edition. Cold Spring Harbor Laboratory Press.

25 Shizuya H., Birren B., Kim U.-J., Mancino V., Slepak T., Tachiiri Y., Simon M.I. (1992) Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. Proc. Natl. Acad. Sci. USA 89: 8794-8797.



Patentansprüche

- 5
1. Nucleinsäure, welche zumindest eine Region umfasst, die für eine Enzymaktivität codiert, welche an der Biosynthese von Spinosynen beteiligt ist.
  2. Nucleinsäure gemäß Anspruch 1, dadurch gekennzeichnet, dass es sich um einzelsträngige oder doppelsträngige DNA oder RNA handelt.
  - 10 3. Nucleinsäure gemäß Anspruch 2, dadurch gekennzeichnet, dass es sich um ein DNA-Fragment handelt.
  4. Nucleinsäure gemäß Anspruch 3, dadurch gekennzeichnet, dass sie alle Regionen umfasst, die für Enzymaktivitäten codieren, welche an der Biosynthese von Spinosynen beteiligt sind.

15

  5. Nucleinsäure gemäß einem der Ansprüche 1 bis 4, dadurch gekennzeichnet, dass es sich um Enzymaktivitäten von Polyketidsynthasen, Methyltransferasen, Glycosyltransferasen, Epimerasen, Aminotransferasen, Dimethyltransferasen, Reduktasen, Dehydratasen und/oder Cyclisierungsenzymen handelt.

20

  6. Nucleinsäure gemäß einem der Ansprüche 1 bis 5, dadurch gekennzeichnet, dass sie aus *Saccharopolyspora spinosa* stammt.

25

  7. Nucleinsäure gemäß Anspruch 1, umfassend zumindest eine Sequenz ausgewählt aus

- 
- (a) den Sequenzen gemäß SEQ ID NOS: 1, 2, 3, 4, 5, 6, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 52 oder 54,
- 30

- 5
- (b) zumindest 14 Basenpaare langen Teilsequenzen der unter (a) definierten Sequenzen,
  - (c) Sequenzen, welche an die unter (a) definierten Sequenzen hybridisieren
  - (d) Sequenzen, welche eine zumindest 70 %ige Identität zu den unter (a) definierten Sequenzen aufweisen,
  - (e) Sequenzen, welche zu den unter (a) definierten Sequenzen-komplementär sind, und
  - (g) Sequenzen, welche aufgrund der Degeneriertheit des genetischen Codes für dieselbe Aminosäuresequenz kodieren wie die unter (a) bis (d) definierten Sequenzen.
- 10
- 15

8. Nucleinsäure gemäß Anspruch 7, dadurch gekennzeichnet, dass sie die Sequenz gemäß SEQ ID NOS: 1 bis 6 umfasst.
- 20
9. Nucleinsäure gemäß Anspruch 7, dadurch gekennzeichnet, dass sie die Sequenz gemäß SEQ ID NO: 4 umfasst.
10. Nucleinsäure gemäß Anspruch 7, dadurch gekennzeichnet, dass sie die Sequenz gemäß SEQ ID NOS: 5 und 6 umfasst.
- 25
11. Nucleinsäure gemäß Anspruch 7, dadurch gekennzeichnet, dass sie zumindest eine Sequenz gemäß SEQ ID NOS: 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37 oder 39 umfasst.
- 30
12. Nucleinsäure gemäß Anspruch 7, dadurch gekennzeichnet, dass sie zumindest eine Sequenz gemäß SEQ ID NOS: 41, 43, 45, 47 oder 49 umfasst.

13. Regulatorische Region, welche die Transkription einer Nukleinsäure gemäß einem der Ansprüche 1 bis 7 in *Saccharopolyspora spinosa* kontrolliert.
- 5 14. DNA-Konstrukt umfassend eine Nucleinsäure gemäß einem der Ansprüche 1 bis 12 und zumindest einen heterologen Promotor.
- 10 15. Vektor umfassend zumindest eine Nucleinsäure gemäß einem der Ansprüche 1 bis 12, eine regulatorische Region gemäß Anspruch 13 oder ein DNA-Konstrukt gemäß Anspruch 14.
- 15 16. Vektor gemäß Anspruch 15, dadurch gekennzeichnet, dass die Nucleinsäure funktionell mit regulatorischen Sequenzen verknüpft ist, welche die Expression der codierenden Regionen der Nucleinsäure in pro- oder eukaryotischen Zellen gewährleisten.
17. Vektor gemäß Anspruch 15 oder 16, dadurch gekennzeichnet, dass es sich um einen BAC-Vektor handelt.
- 20 18. Vektor gemäß Anspruch 17, dadurch gekennzeichnet, dass es sich um einen Vektor handelt, der den BAC-Klonen mit den Hinterlegungsnummern DSM 13010, DSM 13011 oder 13012 entspricht.
- 25 19. Wirtszelle enthaltend eine Nucleinsäure gemäß einem der Ansprüche 1 bis 12, eine regulatorische Region gemäß Anspruch 13, ein DNA-Konstrukt gemäß Anspruch 14 oder einen Vektor gemäß einem der Ansprüche 15 bis 18.
- 
20. Wirtszelle gemäß Anspruch 19, dadurch gekennzeichnet, dass es sich um eine pro- oder eukaryotische Zelle handelt.

21. Wirtszelle gemäß Anspruch 20, dadurch gekennzeichnet, dass die prokaryotische Zelle zur Gruppe der Actinomyceten, bevorzugt zur Gruppe der Streptomyceten gehört.
- 5 22. Wirtszelle gemäß Anspruch 20, dadurch gekennzeichnet, dass die eukaryotische Zelle eine Pflanzenzelle ist.
23. Polypeptid, welches von einer Nucleinsäure gemäß einem der Ansprüche 1 bis 7 codiert wird.
- 10 24. Polypeptid gemäß Anspruch 23, dadurch gekennzeichnet, dass es die Aktivität einer Methyltransferase aufweist.
- 15 25. Polypeptid gemäß Anspruch 24, dadurch gekennzeichnet, dass es die Aminosäuresequenz gemäß SEQ ID NOS: 8, 12, 14, 18 oder 20, oder eine Teilsequenz davon aufweist.
- 20 26. Polypeptid gemäß Anspruch 23, dadurch gekennzeichnet, dass es die Aktivität einer Glycosyltransferase aufweist.
27. Polypeptid gemäß Anspruch 26, dadurch gekennzeichnet, dass es die Aminosäuresequenz gemäß SEQ ID NOS: 10 oder 30, oder eine Teilsequenz davon aufweist.
- 25 28. Polypeptid gemäß Anspruch 23, dadurch gekennzeichnet, dass es die Aktivität eines C-C verknüpfenden Enzyms, das Cyclisierungsreaktionen durchführt, aufweist.
- 
- 30 29. Polypeptid gemäß Anspruch 28, dadurch gekennzeichnet, dass es die Aminosäuresequenz gemäß SEQ ID NO: 16 oder eine Teilsequenz davon aufweist.

30. Polypeptid gemäß Anspruch 23, dadurch gekennzeichnet, dass es die Aktivität eines Enzyms, das an Cyclisierungsreaktionen beteiligt ist, aufweist.

5 31. Polypeptid gemäß Anspruch 30, dadurch gekennzeichnet, dass es die Aminosäuresequenz gemäß SEQ ID NO: 22 oder eine Teilsequenz davon aufweist.

32. Polypeptid gemäß Anspruch 23, dadurch gekennzeichnet, dass es die Aktivität einer 2,3-Reduktase aufweist.

10 33. Polypeptid gemäß Anspruch 32, dadurch gekennzeichnet, dass es die Aminosäuresequenz gemäß SEQ ID NO: 24 oder eine Teilsequenz davon aufweist.

34. Polypeptid gemäß Anspruch 23, dadurch gekennzeichnet, dass es die Aktivität einer 2,3-Dehydratase aufweist.

15 35. Polypeptid gemäß Anspruch 34, dadurch gekennzeichnet, dass es die Aminosäuresequenz gemäß SEQ ID NO: 26 oder eine Teilsequenz davon aufweist.

20 36. Polypeptid gemäß Anspruch 23, dadurch gekennzeichnet, dass es die Aktivität einer Thioesterase aufweist.

37. Polypeptid gemäß Anspruch 36, dadurch gekennzeichnet, dass es die Aminosäuresequenz gemäß SEQ ID NO: 28 oder eine Teilsequenz davon aufweist.

25 38. Polypeptid gemäß Anspruch 23, dadurch gekennzeichnet, dass es die Aktivität einer 3,4-Dehydratase aufweist.

---

39. Polypeptid gemäß Anspruch 38, dadurch gekennzeichnet, dass es die Aminosäuresequenz gemäß SEQ ID NO: 32 oder eine Teilsequenz davon aufweist.

40. Polypeptid gemäß Anspruch 23, dadurch gekennzeichnet, dass es die Aktivität einer 4-Aminotransferase aufweist.

5 41. Polypeptid gemäß Anspruch 40, dadurch gekennzeichnet, dass es die AmiNO:säuresequenz gemäß SEQ ID no: 34 oder eine Teilsequenz davon aufweist.

42. Polypeptid gemäß Anspruch 23, dadurch gekennzeichnet, dass es die Aktivität einer N-Dimethyltransferase aufweist.

10 43. Polypeptid gemäß Anspruch 42, dadurch gekennzeichnet, dass es die AmiNO:säuresequenz gemäß SEQ ID NO: 36 oder eine Teilsequenz davon aufweist.

15 44. Polypeptid gemäß Anspruch 23, dadurch gekennzeichnet, dass es die Aktivität einer 3,4-Reduktase aufweist.

20 45. Polypeptid gemäß Anspruch 44, dadurch gekennzeichnet, dass es die Aminosäuresequenz gemäß SEQ ID NO: 38 oder eine Teilsequenz davon aufweist.

46. Polypeptid gemäß Anspruch 23, dadurch gekennzeichnet, dass es die Aktivität eines Transkriptions-Regulators aufweist.

25 47. Polypeptid gemäß Anspruch 46, dadurch gekennzeichnet, dass es die Aminosäuresequenz gemäß SEQ ID NO: 40 oder eine Teilsequenz davon aufweist.

---

30 48. Polypeptid gemäß Anspruch 23, dadurch gekennzeichnet, dass es die Aktivität einer Polyketidsynthase aufweist.

49. Polypeptid gemäß Anspruch 48, dadurch gekennzeichnet, dass es die Aminosäuresequenz gemäß SEQ ID NOS: 42, 44, 46, 48 oder 50, oder eine Teilsequenz davon aufweist.
- 5 50. Polypeptid gemäß Anspruch 23, dadurch gekennzeichnet, dass es die Aktivität einer Glucose-Dehydratase aufweist.
- 10 51. Polypeptid gemäß Anspruch 50, dadurch gekennzeichnet, dass es die Aminosäuresequenz gemäß SEQ ID NO: 53 aufweist.
52. Polypeptid gemäß Anspruch 23, dadurch gekennzeichnet, dass es die Aktivität einer 3,5-Epimerase aufweist.
- 15 53. Polypeptid gemäß Anspruch 52, dadurch gekennzeichnet, dass es die Aminosäuresequenz gemäß SEQ ID NO: 55 aufweist.
- 20 54. Enzyme, die an Cyclisierungsreaktionen beteiligt sind, dadurch gekennzeichnet, dass sie die Aminosäuresequenz gemäß SEQ ID NO: 15 oder 22, oder eine Teilsequenz davon, welche zumindest noch eine Teilreaktion durchführen kann, umfassen oder eine mindestens 50 %ige Identität dazu auf Aminosäureebene aufweisen.
- 25 55. Antikörper, welcher spezifisch mit einem Polypeptid gemäß einem der Ansprüche 23 bis 54 reagiert.
56. Verfahren zur Herstellung einer Nucleinsäure gemäß einem der Ansprüche 1 bis 7, umfassend die folgenden Schritte:
- 

- (a) Vollständige chemische Synthese auf an sich bekannte Weise oder

5 (b) chemische Synthese von Oligonucleotiden, Markieren der Oligonucleotide, Hybridisieren der Oligonucleotide an DNA einer genomischen oder cDNA-Bank, die ausgehend von genomischer DNA bzw. mRNA aus *Saccharopolyspora spinosa* hergestellt wurde, Selektieren von positiven Klonen und Isolieren der hybridisierenden DNA aus positiven Klonen oder

(c) chemische Synthese von Oligonucleotiden und Amplifizierung der Ziel-DNA mittels PCR.

10

57. Verfahren zur Herstellung eines Polypeptids gemäß einem der Ansprüche 23 bis 54, umfassend

15

(a) das Kultivieren einer Wirtszelle gemäß einem der Ansprüche 19 bis 22 unter Bedingungen, die die Expression der Nucleinsäure gemäß einem der Ansprüche 1 bis 7 gewährleisten, oder

20

(a1) das Exprimieren einer Nucleinsäure gemäß einem der Ansprüche 1 bis 7 in einem *in vitro*-System, und

(b) das Gewinnen des Polypeptids aus der Zelle, dem Kulturmedium oder dem *in vitro*-System.

25

58. Verfahren zum Herstellen von Spinosyn-Derivaten, einschließlich Spinosyn-Vorstufen, umfassend die folgenden Schritte:

(a) Austauschen zumindest einer Modul-codierenden Nucleinsäuresequenz gemäß Anspruch 7 gegen zumindest eine andere Modul-codierende Nucleinsäuresequenz gemäß Anspruch 7, oder

30



- (b) Austauschen zumindest einer Modul-codierenden Nucleinsäuresequenz gemäß Anspruch 7 gegen zumindest eine andere Modul-codierende Nucleinsäuresequenz aus *S. spinosa*, oder
- 5 (c) Austauschen zumindest einer Modul-codierenden Nucleinsäuresequenz gemäß Anspruch 7 gegen zumindest eine andere Modul-codierende Nucleinsäuresequenz aus einem anderen Organismus als *S. spinosa*, oder
- 10 (d) Austauschen zumindest einer Domänen-codierenden Nucleinsäuresequenz gemäß Anspruch 7 gegen zumindest eine andere Domänen-codierende Nucleinsäuresequenz gemäß Anspruch 7, oder
- 15 (e) Austauschen zumindest einer Domänen-codierenden Nucleinsäuresequenz gemäß Anspruch 7 gegen zumindest eine andere Domänen-codierende Nucleinsäuresequenz aus *S. spinosa*, oder
- 20 (f) Austauschen zumindest einer Domänen-codierenden Nucleinsäuresequenz gemäß Anspruch 7 gegen zumindest eine andere Domänen-codierende Nucleinsäuresequenz aus einem anderen Organismus als *S. spinosa*, oder
- 25 (g) Austauschen einer ersten Acyltransferase-codierenden Nucleinsäuresequenz gemäß Anspruch 7 gegen eine zweite Acyltransferase-codierende Nucleinsäuresequenz gemäß Anspruch 7, wobei die zweite Acyltransferase eine unterschiedliche Substratspezifität aufweist als die erste Acyltransferase, oder
- 
- 30 (h) Austauschen einer ersten Acyltransferase-codierenden Nucleinsäuresequenz gemäß Anspruch 7 gegen eine zweite Acyltransferase-codierende Nucleinsäuresequenz aus *S. spinosa*, wobei die zweite

Acyltransferase eine unterschiedliche Substratspezifität aufweist als die erste Acyltransferase, oder

5

- (i) Austauschen einer ersten Acyltransferase-codierenden Nucleinsäuresequenz gemäß Anspruch 7 gegen eine zweite Acyltransferase-codierende Nucleinsäuresequenz aus einem anderen Organismus als *S. spinosa*, wobei die zweite Acyltransferase eine unterschiedliche Substratspezifität aufweist als die erste Acyltransferase, oder

10

- (j) Deletieren zumindest einer Domänen-codierender Nucleinsäuresequenz gemäß Anspruch 7, oder

15

- (k) Integrieren zumindest einer Domänen-codierenden Nucleinsäuresequenz gemäß Anspruch 7 in eine Modul-codierende Nucleinsäuresequenz gemäß Anspruch 7, oder

- (l) Mutagenisieren zumindest einer Domänen-codierenden Nucleinsäuresequenz gemäß Anspruch 7,

20

und Exprimieren der rekombinierten Nucleinsäuresequenz in einer Wirtszelle unter Bedingungen, welche die Synthese eines Spinosyn-Derivates oder einer Spinosyn-Vorstufe erlauben.

25

59. Verwendung einer Nucleinsäure gemäß einem der Ansprüche 1 bis 7 zum Identifizieren, Inaktivieren und/oder Modifizieren von Genen der Spinosyn-Biosynthese.

30

60. Verwendung einer Nucleinsäure gemäß einem der Ansprüche 1 bis 7 zum Erzeugen einer Bibliothek aus Polyketidsynthasen.

61. Verfahren zum Anfügen eines Forosamin-Zuckerrestes an das Spinosyn-Aglycon oder das Spinosyn-17-Pseudoaglycon, umfassend die folgenden Schritte:

- 5
- (a) Übertragen einer Nucleinsäure gemäß SEQ ID NOS: 23, 25, 29, 31, 33, 35 und 37 in eine Wirtszelle,
  - (b) Zufügen des Spinosyn-Aglycons oder Spinosyn-17-Pseudoaglycons, und
  - (c) Kultivieren unter Bedingungen, die zu einem aktiven Stoffwechsel der Zelle führen.
- 10

62. Verfahren zum Anfügen eines Trimethyl-Rhamnose-Zuckerrestes an das Spinosyn-Aglycon, umfassend die folgenden Schritte:

- 15
- (a) Übertragen einer Nucleinsäure gemäß SEQ ID NO: 7, 9, 11, 13, 17 und/oder 19 in eine Wirtszelle,
  - (b) Zufügen des Spinosyn-Aglycons oder Spinosyn-17-Pseudoaglycons, und
  - (c) Kultivieren unter Bedingungen, die zu einem aktiven Stoffwechsel der Zelle führen.
- 20

63. Verfahren gemäß Anspruch 62, dadurch gekennzeichnet, dass im Schritt (a) Nucleinsäuren gemäß SEQ ID NOS: 9, 11, 13 und 17 übertragen werden.

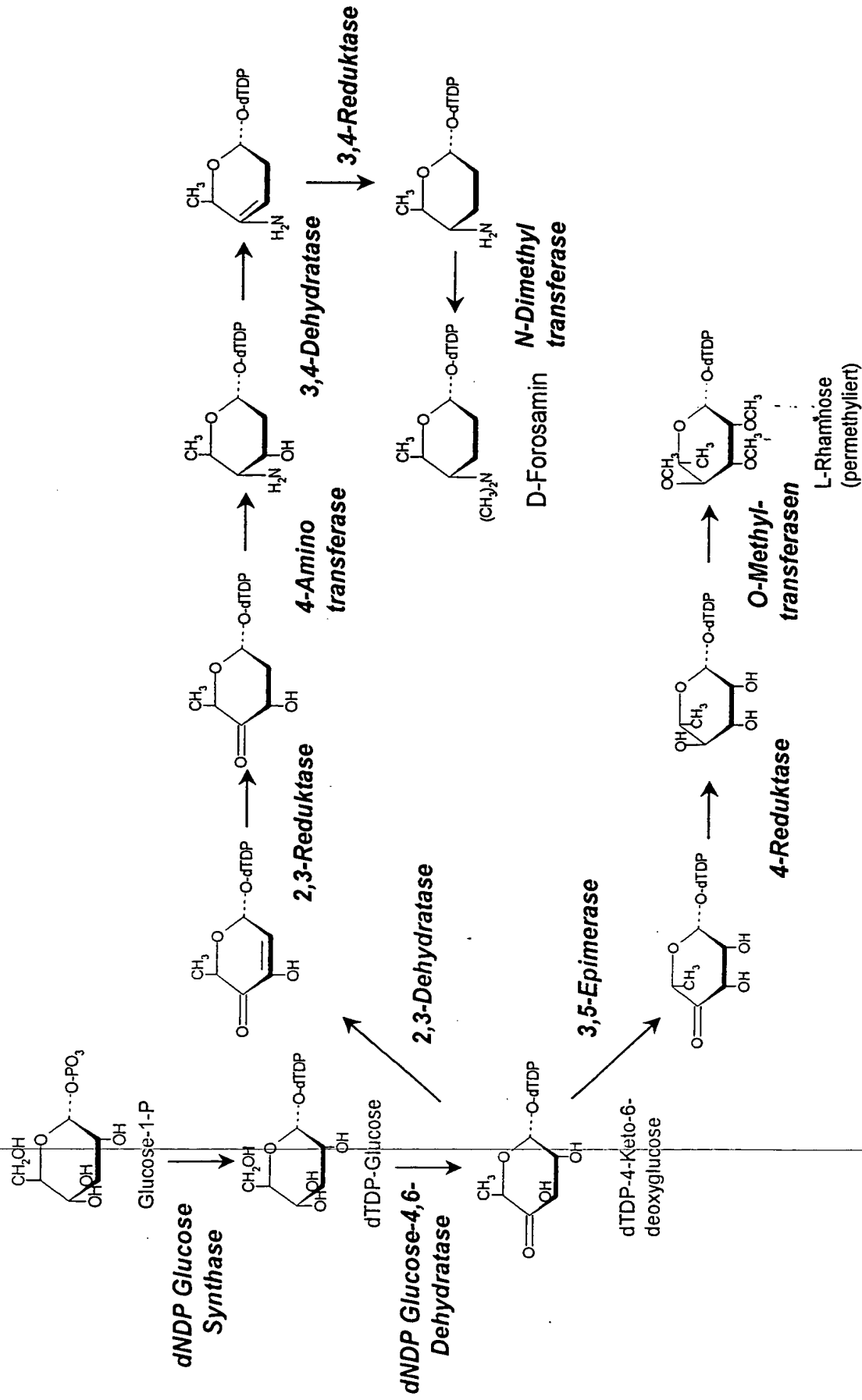
---

Nucleinsäuren, die für Enzymaktivitäten der Spinosyn-Biosynthese codieren

Z u s a m m e n f a s s u n g

Die vorliegende Erfindung betrifft Nucleinsäuren, die für Enzymaktivitäten der Spinosyn-Biosynthese codieren, sowie die entsprechenden Enzyme per se. Weiterhin betrifft die Erfindung Verfahren zum Herstellen von Spinosyn-Derivaten und -Vorstufen.

**Abbildung 1: Metabolismus für die Biosynthese des D-Forosamins und 2, 3, 4 Tri-O-Methyl-L-Rhamnose**



**Abbildung 2:** Lage, der Spinosyn Biosynthesepfad oder indirekt beteiligten DNA-Regionen 1 und 2 sowie Lage und Bezeichnung der zur Sequenzierung herangezogenen überlappenden Cosmidklone

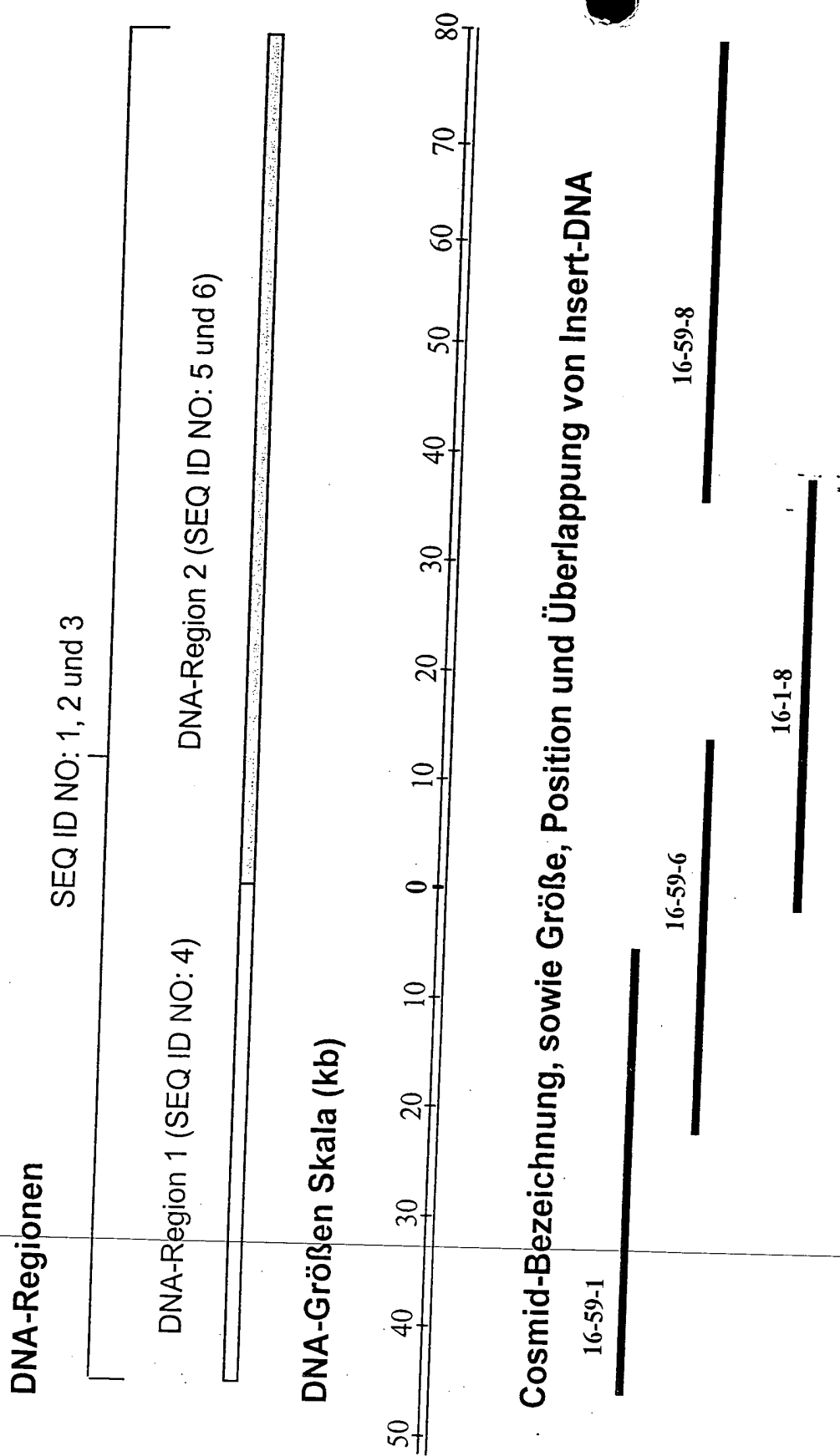


Abbildung 3: **Electroporation herangeführte Cosmidklone**

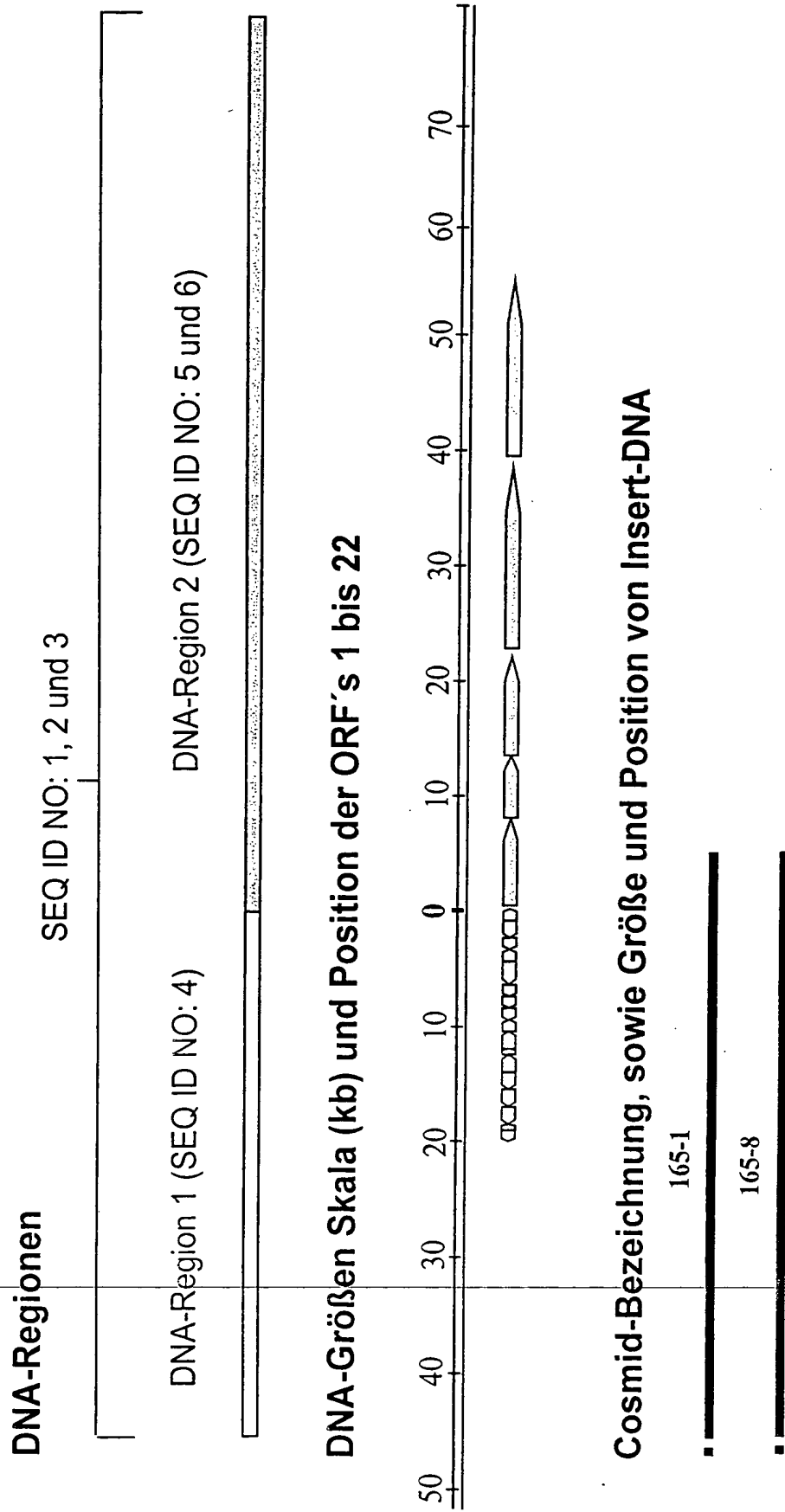
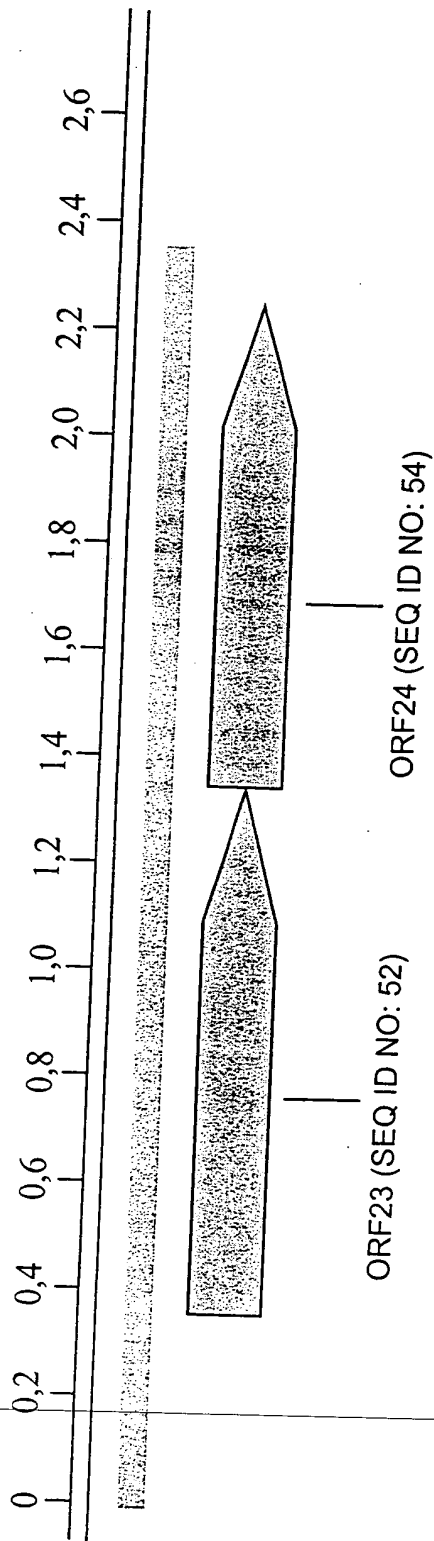


Abbildung 4: Darstellung offener Leserahmen  
der DNA-Region 3 (SEQ ID NO: 51) auf Cosmid 16-2-2

SEQ ID NO: 51

DNA-Größen Skala (kb) und Nucleotidposition





**Abbildung 5:** Darstellung offener Leserahmen der DNA-Regionen 1 und 2, die direkt oder indirekt an der Spinosyn-Biosynthese involviert sind

## DNA-Region

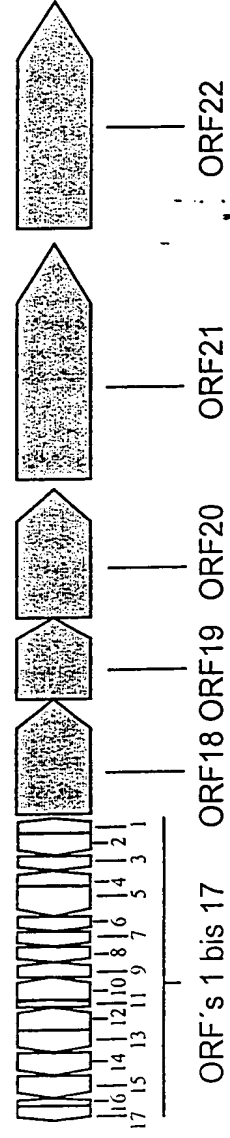
DNA-Region 1 (Ausschnitt)

DNA-Region 2 (Ausschnitt)

## DNA-Größen Skala (kb) und Nucleotidposition



## Offene Leserahmen (ORF's)



Stellung offener LeserInnen

SM	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10
KS AT ACP KS AT DH	KS AT KR ACP KS AT DH	KR ACP KS AT DH	KR ACP KS AT KR ACP KS AT DH	KR ACP KS AT DH	KR ACP KS AT DH	KR ACP KS AT KR ACP KS AT DH	KR ACP KS AT KR ACP KS AT DH	KR ACP KS AT DH	KR ACP KS AT DH	KR ACP KS AT DH

**Abbildung 7: Darstellung von Spinosyn Biosynthesegenen BAC Shuttleklonen**

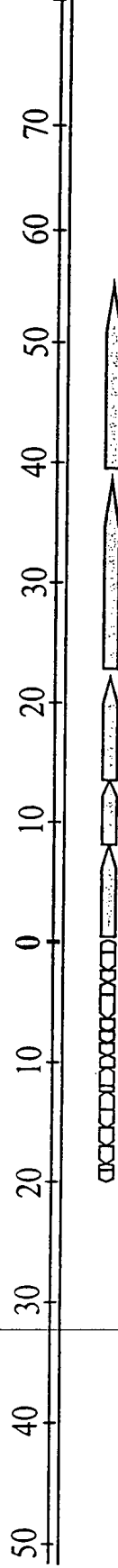
**DNA-Regionen**

SEQ ID NO: 1, 2 und 3

DNA-Region 1 (SEQ ID NO: 4)

DNA-Region 2 (SEQ ID NO: 5 und 6)

**DNA-Größen Skala (kb) und Position der ORF's 1 bis 22**



**BAC Shuttleklon-Bezeichnung, sowie Größe und Position von Insert-DNA**

P11 / G6 (> 100 kb)

P11 / B10 (> 100 kb)

P8 / G11 (> 100 kb)

# SEQUENZPROTOKOLL

<110> Bayer Aktiengesellschaft

<120> Nucleinsäuren, die für Enzymaktivitäten der  
Spinosyn-Biosynthese codieren

<130> Le A 33 955

<140>

<141>

<160> 55

<170> PatentIn Ver. 2.1

<210> 1

<211> 50000

<212> DNA

<213> Saccharopolyspora spinosa

<400> 1

```

ggcaccaccc cggggccccc gcctcactcg gagtcgatcc gcattccggt cacgaacatg 60
tgaccgcgtcg agtgcgatg cattagctcg gtcttcgcct cgatcgccgc ctgctgcggg 120
gtcacgccgc agggccagaa caccggcacg tcgcccggct ccatacgcac tcgttcaccc 180
cagtcagggtt catcaacgtt gtgatgccc aatttctcag gggagccgat gtgcacgggg 240
ccccgtgca cctcgggatg cgcgtgctg atggtgacag cctccgccac ccggtccggg 300
ggcatgggac gcattgaaac cactagcttc ccccggaact ggctgcggg cacacactca 360
cggttgctga ggtaaaccgg gccgttgatc cctgggtcga tgtgccggat cgggagtccg 420
gcatcgagca gcaagtgtc gaaggagaag ctgcagccaa gcaggaacgc gaccatgtcg 480
tcggtccaca ggtcaacgac gtcactgcgc tcctccacca gcgccccttc ccggtacacg 540
ctgtatttgg gcaggtcggg gcggagggtc gcaccagggg cggtcagctt cggctccgga 600
tctccgcgt cgggtgacct gaggatcggg cacggcttgg ggttcgctg ggccaaccgc 660
atgaagtcga acgcgtcatc ctggggcagc atcacgaggt tggcctgcac gaatccatga 720
cagatcccg acgtgggcct cggccacgca ttgcgcggt tggcacgcca tgcctcgatg 780
ggctccgcgt gctgcagatc aacgaactcc atcgatgact tccccttccc ggcgatgatg 840
acgaaggctg gcaaccgagg gttgcccgct acgggtggcc gtcactggtc tgccacccag 900
gggtcgccga gcaggctctg gtgaacgatt ccggagatca gtcggcatg cgcggcgacg 960
aaacgcagtg ggatcccgtc gtgccttgg aacgcgggat cgtcaacgcc gccgtgcgt 1020
agacctgcgc gcagccccag ctgggtccac tcgcgcacca gatccccggt gactgccagg 1080
tttgcgctc gagccaggat ggccaccgca ccggtgatga cgtagcagcc caaattcgac 1140
acggatgcag gcagcagatg atccaccgtc gtggacgca agaattccgc ctcaacgctg 1200
acgcctcgcg gagtcagcgc ctggactcg gcggcaatgt tgccgaaccc gatctcgttg 1260
ccgttgctc ccacggccag cgtcagagca ccgttgctcc tggcggcgcg aaggtagtcg 1320
tcggcaaaca gatcaccggc ggtgaccggc gtgcccaga tcagatgacg gccgccaaca 1380
gagttctgcc ccagcttctc gatggcggcg acgacgccga actgcgccgc gaactcccg 1440
gcctcggccg cgctggagaa tcgggactcc atcaagggtc cgcctccag gccgcagcg 1500
tcgaccagct tctccaccgg gccgacggcc tctgggtcga tgacgatgg cacctcgcta 1560

```

ccaagaccgg	cgagcgtg	cgcgagtgc	agagagccga	tggggccg	cacctcccc	1620
tgcgggaacc	ggccttcctc	gaccagaccc	gtgacgatca	acgccttttc	accctgtcga	1680
cgcagcagag	agcgcgcgat	cgcaaggctg	atcggtccg	accgctcacc	ccgcgccagc	1740
gtgtagaacg	ggctcacgat	gccggccggg	ctcgtcggcg	gacgcattctc	gatcgtggcc	1800
agtcgatcca	ggctctcgaa	ggccgggtgc	gactcgggct	tcattcagctc	tcctcgggta	1860
tacgacatat	cccaaggact	tcattccgcta	cggagtcccg	cctcggggcg	ttgcttacgc	1920
ggtgaggacc	atataccgta	ttccactgtc	caaggaagg	ggacgatgga	accgaatcag	1980
ccacgcactg	cgggcacgct	gaacatggtg	gagctgcacc	atcgcccagc	cgccgcagac	2040
tccgcggtgg	ggtccggcaa	cggtatgctc	gcacccctcg	acgcacgcgt	cgacaattctg	2100
acagatcggg	tcttcgaggc	catccgcgac	agcatcgtaa	gtcaggcact	cccacctggg	2160
cgtcgtgtga	gcgaggccaa	gctcgcggcc	gagctgaatg	tcagcaagac	ccccgtgcgc	2220
gaggcgcttc	tccgcttg	gcacgtgggc	ctggtcgagg	ctgccccgaa	ggggctgcgc	2280
gtcatcagcg	cctcgggtcg	cgccattcgc	gacgcctacg	agcaccgggc	ggggctggag	2340
agctcggcag	cctgggttcac	cgcacatcgg	gctaccagcg	cccaacggga	gaggctccac	2400
gaggtcgcca	cctcatccct	gcgcgcggca	gagtcggcg	actccgccc	gttcgcgcgc	2460
gaggatcgcg	tcttcaccca	catcatcgcc	gagttctg	ggaacgagat	cctccagcgc	2520
gcggtcgaca	acgcctggt	gctgaccgct	gtactccgag	aacgcgacac	tccttcgacc	2580
ggtgactcca	tctcctgcgc	ctcggagcac	atcgcgacgg	catccgccgt	tctggctggc	2640
gacgccgacc	gggcagcccc	gaacaacgcc	gaccacatcc	tccacgtgat	gtcgttgtc	2700
ctctcctcac	acacgcagcg	cggaacgtg	cccccgga	cgagttaa	caactcccc	2760
acccctacc	tcaccattcg	accgaaggac	gggtttcgt	gaccgagcac	ctggcgacca	2820
tcaccgctgc	ggaaggagag	atccccgttg	tcgtgagcga	ccgcggcg	gcagcaatcc	2880
gggagataga	ccccgagctg	gcggtcacga	ccatcatgga	cgtgatcgca	ggagaccacc	2940
tcgcgcggct	gcgtgagcgt	gtcaagaacg	ccccgacct	cgcttcaca	cgctggacg	3000
aggtgcggtt	caccgcgcgc	taccgatccc	cccggaagat	ctggggcatc	ggtctgaact	3060
acggcgatca	cgcagcggac	ctgagcaggt	cggcgcccga	ccagccggcg	tccttcgtga	3120
agtgcgacca	cacgatcatc	ggcccggatg	agccgatcgt	gattccggcg	cagagtgage	3180
gcaccacgtc	cgaggctgag	atcgggctca	tcattcggcag	gacttgcgag	aacgtcagcg	3240
agtccgaggc	gatggactac	gtgtggggg	tcacgaccat	cctggaccag	accgccgagg	3300
acatcttg	gatcaatccg	cgctatctca	ccgcctcgaa	gaacttcgg	acgttcttct	3360
gcttcggacc	ggagctcgta	ccgctcgatg	aggtgataga	gaagttcgaa	agcttggacg	3420
acatcgagat	ctcgacggtc	aagaacggcg	atgagttcag	gaccaacacg	gtcgcccaca	3480
tgaccacaa	ggccgccagc	ctcgtgagct	tccactcgaa	gatgatgccg	ctgtttccgg	3540
gagacatcat	ctccaccgga	acccctggcg	cgctcgtcat	cggcgacgga	gacgtgtgg	3600
agtgcgggat	cgacggcatc	ggtcagctga	ccacgacggc	gcgaaaggcc	actacacgat	3660
gagcactcac	gcggtcgagc	agctcgaccg	gaccgcactc	ctgcaatgtg	ccgacgacct	3720
ccgcgagacc	gtcctcgacg	acgcactcaa	ggcggtgcc	ctgggtgctc	gtgtgctcct	3780
cggagagggtg	gaccagcaga	actggaacgt	ggcccgcgga	gacctcgccc	tgccgggtgac	3840
cacactccac	gcacgcgcta	tgaacgcaa	cctggcgact	atggaggcgt	actgcgagcg	3900
caacaacgcg	cttctcgccc	cgcacggcaa	gaccaccatg	tcaccccagc	tcttcgccc	3960
tcaggcgtct	cacggtgcgt	ggcggttcac	ggcagcgacg	ccactcagg	tggccgtgat	4020
gcgcgcgtac	ggggttcagc	gcattcatcct	ggccaacgaa	ctcgtcgatg	ctccgcgcgt	4080
gcgttggttg	actgccgaaa	tcgccagcga	cagctcgttc	gacttctact	gcctggtcga	4140
cgatgcggac	acggtgcggg	cgatgtccga	taccatcgcc	gactccggcc	acggggtcac	4200
cctcaacgtc	ctcctcgagg	tcggagtgc	tggtggcagg	tgtggcggtc	gagacctcga	4260
ctcggcgctc	gaggtcgccg	cggcggtcgc	cgcgaccgag	catctccggc	tggtcggcgt	4320
cgaggcctac	gagggcctcg	tgaccggcgg	catgacaccg	gaggacctcc	atgcaactcga	4380
caccttcttc	gctggcgctc	ggtcgatcgt	gctcgagctg	gcacgtcgcg	acctgttcga	4440

cgccaaccgc atcatcgga cggcaggcgg aagctcgtac ttcgatcgtg tctcgcgccg 4500  
 actcggctcg tgggacggtg tccgcgacga cgtggaccta gtgcttcgaa gcggctgtta 4560  
 catctccac gacgccggca agtacgaaaa gctctcgccg ctcgctggcc gacgggccga 4620  
 gcaggaaccg ctccggctcc acaacgcact cactgcatgg gcttcctgctc tgtcccgcc 4680  
 ggagcctgat ctgcgatcc tgcgatcgg gaagcgcgac gcggcccatg acctgacgct 4740  
 gccacagcca cgcgagctgt ttgcgcgga tggctcgagg cagcggctgc gggaagcgg 4800  
 gacgttcaag ctgatggacc agcacgcctt cgttcgcggt gcccccgggc tgcgatcgc 4860  
 tcccggagac atcgtcgtt tgcacatgtc ccaccctgac acggccttcg acaagctccc 4920  
 ctcatcccg atcatcgacg acgacttcaa cgtcgtcgac ggcgtcctca cgttctctc 4980  
 acggggctca ggacctccta cgggaggctc cgggtgaacc gacccacaa gctcccgaag 5040  
 tcagttggac atgaccacca gccacagaga ggccccgaca ccgtgtcccg tcaggcgttc 5100  
 acagacgtcg accttctcc cgcgggaggc ccttactccg ccgccgtacg catcggcaac 5160  
 atcgtggcaa tgcgcgggca gtgcggctac ctcgccgaca gatccctcgc accgggtggt 5220  
 ctggaacctc aggtgcgccc ggcattcgag aacctcatga ccgcgctgcg ttcagccggc 5280  
 tgcagcgagg cagacgtgat caccgtgaat gtgttctcgt ctgacgggga cgacttcgac 5340  
 gcgatgaacg cgatctaccg cgagtacttc agcgagcccc gcccgccccg gacgaccatc 5400  
 accgccggtc tccgaccggg tgcctgttc gaggtcagcg ctcaggccgt ggtaggcggc 5460  
 tgaccgggccc actgccactc ctcccaccaa tgccttcca tcaacggcgt gccacgactt 5520  
 tccgggcccgc cgtcacccga aaccgcaagg acatgtcatg gatctcgcaa tcaatggaag 5580  
 agtcgccctc gtactcggag ccggcgggcg actcggttcg gccattgccg cctcgttaag 5640  
 ccgggaaggc gcgcgtgtgg cgctggccga catcgacaag gacgcgtca cggaggctga 5700  
 ggcgtccatg accggcgata acatgagcct ggtctgggac ttggcgaacc tgcgggtcat 5760  
 cgatgagaac gtcaccgccg tcgagagcga gtcggccccg gtggagatcc tggtaacat 5820  
 caccggcggc ccgcgcacca cccgtgtgtc cgggcaggac gccaccctgt ggcgtcggca 5880  
 cttcgaggcg atgatcctc cgtcgtcgc gatcacagac cgcgtgctcc cgggcatgac 5940  
 cgagcgggga tggggccgga tcatcacgtc gacttcgtcc ggcgtcatcg cgccatccc 6000  
 caacctcggg ttgtcgaacg ccctccgctc gagtctggtg ggctgggtcca agacactgtc 6060  
 cggtgagggt gcaagcgacg gcatcaccag caacatcgtc gttccggggc gcatcgccac 6120  
 tgcgcgcatc gccagttgg accaggcaaa ggccgcacgg gagaaccgca ctgcggagaa 6180  
 catcgcgagg gaatcagaga gctctattcc ggtcggccgc tacgggcccgc gcgaggaatt 6240  
 cgccgacgcg gtcacatttc tcgccagcga cgtgcccagc tacatcaccg gatccgttgt 6300  
 ccgtgtcgat ggcggactgc tccagaacgt gtaaccccga caaacatcct gtctccagat 6360  
 cagcgcaacc cgaagccatg tccgtcctc ggggtcactg acgtcagcga tcccagccc 6420  
 aaactcatca cctacgtgcg ctgatctacg gtcattcagg gctaccact cgccagcggg 6480  
 atcacgatca gcagcacgga cctcgctcc aggggtgcag ctgcgcgcga gcggtgccgt 6540  
 cggttttcgg gtgcgcggcg accagcttt tccccacatt gaaagctatc cggaaccgga 6600  
 tgctgcctgg aaggcgtact tcaggtatgg atgccgaac gcgaccttca gctacctacg 6660  
 atcgtatctc tggaaagagg tgcctcgtg gcaggagcgc aaacaccggc gtactccctg 6720  
 gaaacagcta cgcgacgct acggcatcca gccagccgac ggcgatgtcg ggctgttcga 6780  
 tccggaaggc cggctcggga gacgtgcagc cgggtgggga aggagcgtg acaaaatcct 6840  
 gacgacctgg cactaccagg acagctcgtg ggcaagcgag atcaggcggg gtatgcgaaa 6900  
~~cctcagatct tgetaacagt caageggaga cccgttctat ggtcgtgtta tctcgcggc 6960~~  
 caccgcgct ggctcacgt ctgtgcgcca gcggaatgt cgcacaggcg cgtattccc 7020  
 gaaaggaacc catcgaatca tgacgaaccg tcccttcgc tggctcgtcg tggcgccag 7080  
 cgacatcgcc cgcgcatttg tagggccagc gatctccgc cagcctgacg ccgcggtggt 7140  
 cagcgtcgtc agcggcgacc tggcccgagc gcaggatgta gcggcaccat tcggcgcccc 7200  
 gggccatgac gaactaccg tggcgcttag cgtacctgac gtcgacgccg tgtacatctc 7260  
 gagtatcaac tcggcacatc acgatcaagc acttgccggc atcgcagccg gcaagcacgt 7320

cctgtgcgaa	aaaccgcttg	ccctgaccgt	cgccgaggca	caagaacttg	tcgatgccgc	7380
ggagaaggca	ggcgtagtgt	tcgcgacaaa	ccaccacatg	cgcaactccg	tgccacatcg	7440
tttgatccgg	gatgcaatcg	ccgcccggcg	catcgggcag	ccggtcgcag	tgaacgtccg	7500
gaacgcgata	cgactgccgc	ccgcccgcgc	gagatggcgg	acgaccgatt	cggcagcggg	7560
cgcccggagtc	gcactcgacc	tcacggtgca	cgacgcggac	tgtctccggt	tcgtactcaa	7620
tgccgatccg	ctaacggtcg	tggcccatac	ctcgtccggc	cacatgacga	gggtgggaat	7680
cgatgagaca	atctccggaa	gcgcgaggtt	cgtcagcggg	gtgcatgctt	cattcgtgga	7740
aagcttcgtg	accggacatg	ccccacatg	cctggaggtt	ttcggaaccg	aaggagcgct	7800
catcggcacg	ggtatccagt	cgatgtcacc	ggtcggcacg	ctgacacacg	tcggtccggc	7860
cggtccagg	aagatcgatc	tgggtaccgc	cgaggacctc	tacgtcgtgg	gtgttcgacg	7920
gttcatacgag	gctgtgcgtg	gcgaggccga	gcccccgcc	tccggccaag	acggggtctg	7980
gtcgttggcc	ttcgacacgg	cggcgctgga	ggctgctcgg	accgggcatc	ggcagactgt	8040
gggacaggct	cgatgacgaa	agctgggcgc	agttaggcac	aacagatcat	gcacaaggac	8100
aaccgggcat	cggcagacgt	ggcgatcgtc	gggtccggtg	gaggctgact	gatcgtcgcc	8160
ggggcgctcg	ctggtcgcgg	actcgatgtg	ctcgtgttgg	aacgggacga	gtttttggct	8220
caagagcccc	ctatcgacgt	gtgcaaggct	tcgcgcacgg	gcactacaag	ccggaögaac	8280
gggtggctcga	tggtgctggg	agaccgttca	caccggctac	tcactactcg	tcggcggcaa	8340
cagccaaggc	ctacgaagct	ggcccgtggc	tcaatctata	ccgcgtaagc	tcgccaacgg	8400
gtcctatgtg	ggcgcccgca	actacatgca	gcacccgaac	tccgcggtca	tggccatcgg	8460
caatggccag	cctccgatca	tggaccatac	ctacttcaac	atgcgccggt	gcggggggccg	8520
cggacgatta	gctgaagcgg	atgtccgctt	tcggtatcgc	agaccttgcc	acgacggggc	8580
gtgttgccgg	gccgcgaacg	gatcatgtga	gccgggctgg	cccagagaagg	gcatgggcca	8640
acagagcagg	tgactacgtt	atctctgccc	ggatcatccaa	ctgatgacct	gtgcgcgcga	8700
ggtaaagccc	agcttggaag	ggatgttctc	gacgtgggcc	tcggctgtgc	gctgggaaat	8760
gaccaggcca	gccgcaatct	ctttgtttgt	tttgccagcg	gtgaccaggg	cagcgacctc	8820
gcgttctcgt	tttgtgagaa	ccggttcggt	ttccggagcc	gaagcggcgg	ccgcttcgac	8880
cggcgggtgg	ttctcccca	gagcatagaa	gatagtctgc	tcccggttca	tccgcgtccc	8940
ctggcctcgg	gcggtttcgt	attcgggtgc	tgtgagtgtc	tcctgggcgc	gtgccttggc	9000
ctgctcacgc	caggtcagta	gcgtctcgga	tccgaacagc	gggtgcca	tcagttcca	9060
cattttatcc	aacgccccga	aaaggatcgc	cgcgcgttca	ggattgcgtt	ctgcggcggc	9120
gacccaggcc	agcagttcaa	cgcagtcgct	gatcccaggt	aggctcgttg	gttcccgtt	9180
attgcggagg	gacctggacg	cggaggcggc	tgccttggtt	gggtcgccc	cggcccacca	9240
agtgatcgcc	atgttccact	ccgcccacga	tagcctccag	cgctcgccaa	gttctgtaca	9300
atcttgcgg	cattcgtgca	gaaggacat	cgcccgggtc	aggtctccaa	gcaggcccgc	9360
tgccctgcgt	cgtattgaga	aaatggctaa	ggcggggccg	gtccagtggc	cagttgcgcg	9420
gtggcgggcc	agtgcctcgt	cgagcagttg	cgtggcgcg	gtcaggttgt	tgtccacat	9480
ctcccggtcg	ccaagaaact	gagtcgcgta	ggtaaattcc	gtttcgtcgc	ccagtttccg	9540
ggccagctcg	cggctttcgt	taagcaggac	cagcgcccca	tctcgggtcc	cttgccggtg	9600
ggcaatccat	ccggtgatcc	acagtgcgcg	ggcacgctgg	ctggtgggtt	cggctctccag	9660
ggccaacgcc	ctgtccagcc	acatgcggcc	ctccctcaga	tagtcgccag	cagcccagta	9720
gaaccacagc	gcgctcacca	tccgcaaccc	aaccgacgcc	tcgcccaggt	tgggtgaagca	9780
<del>gtagteeeaa</del>	<del>getgeccaga</del>	<del>agttgeegcg</del>	<del>cteggegege</del>	<del>agaeggegga</del>	<del>geaaeeegge</del>	<del>9840</del>
ttgatgcggg	cccgcggact	cggcatcgga	ctgctcggcg	agatgcaggt	agtagtcgcg	9900
gtgccggcag	cgcagagcgt	tttcggcacc	cgcctcgggt	agccggtcgt	gcccgtattg	9960
acggatagtc	tccagcatcc	ggtatctcgt	cttcgtgcgc	tcctcgggtc	gggtgagcac	10020
ggacttgtcg	atcagtcccc	cgagagcggc	gaagacgtca	tcacgtgtca	gcccgtcgcc	10080
ggtgcagacg	ttttcggtcg	cgtcgagatc	gaattcgtcg	gtgaaaaccg	aacagcgcgc	10140
ccacaactgc	ttctcatgct	cattgcacag	gttgtagctc	cagtcgacgg	cggcgcgcg	10200

ggtctgatgc cgaggatgc cggcgcggtc gccgggtggtg agcagcgcga atcggttctc 10260  
 cattcgcggtt agaatttgcc cgactgagag cagcogaacg cgcacggcgg cgagttcgat 10320  
 cgctagggga atgccgtcca gcctccggca taacctggcc acgatctggt cgttttcttc 10380  
 gttgatggtg aaccttggtg gcgcgcgcgc ggcacgctgc tcgaacaagg tcaacgcttc 10440  
 acgctggtag ggacgacctt cctggttgta cgttgatgaa tcgctgccag caggcacgga 10500  
 cagtggcgga accggccagg tctgttccgc cactatccct aaatgttctc ggctggttgc 10560  
 aagtatccgc aaccgcgggg ccgcggccag cagcacagcg actaggcggc tgcaggcccc 10620  
 cagcaggtgc tcgcagttgt ccaacaccag caagacttgc ttgtccgcca agtaatcgac 10680  
 gagcacggcc tccggggcac gtgtcgaacg gtcattggatc ccaagagcag ccgcgaccgc 10740  
 ctgcggcacc agcatgggat cgcgaacgtt agccagctcc accagaaatg tccggcgga 10800  
 gaaggccccgc cgcacgtcgt gtgcgacgtg cagcgccagc cgggacttcc cgatccccgc 10860  
 gaaaccagtc aggggtcacta accgtgaact ggacagcgcg cgcttaactc ttgtggccgc 10920  
 ctgacggcgc cccacaaagc tcgtcacctc atatgggagg cgaccagcct cccgaaccgc 10980  
 aaccccagac acgactcacc gcacctcggg agaactagcg ttttattgta gcccggtcat 11040  
 ggctggcagc gcacttacag cttgaacaac tggtcgaacc acttcgaacc caccgacggg 11100  
 gttcgttacg atctcggccg atgttgatcg tgggtgggtgc ggggagggcg tgcgttgccc 11160  
 gctggcctgc ccagcctgta ccactgatca aggtcgccga cggggcccggt gtgggcgatg 11220  
 ttggcactgg tcaggccact gctggcggtt atgcggcggg ccagctgcgg cggcacggtg 11280  
 ggcggcatcc gtcctgctac tgggtgcgca aggcgctcgc gtccgcctga cggcgaaagg 11340  
 tgagctcgtc gtggaggcag cggaccgggt gatcgctgag ctggagacgc tgcgactca 11400  
 gatcaatcgc tcacgtggcg agatctcgca acgtctcgt gtggcctcca tgcagaccgt 11460  
 caccttcac cgttccccg aggtgctgac tcgtctacgc cggcggcgtc caggactgga 11520  
 tgtggtgatt gttcaagccg agccggatgc ggctcttcag ggtctccagg tcgacgacta 11580  
 cgacctagt atcgccgagt cctaccgcga gtggccggtc atctatccgg acgggctcac 11640  
 gacagagatc ctgatgaccg accgatgca cgtcgccgca gctcccattg tgggtccacc 11700  
 gcccgctcgc tgggggcagg cgctttcgcc gctggcgagt cgttcaccgc tctcggcgct 11760  
 ccaacgtctt gcctgctcga agttgtcccg agtcgggccc gacgatggct gtggtgaagg 11820  
 gatcttgacg ctcatgcatg gaggttcacc attgttgagg caagaggctg atgacggtg 11880  
 cggcggtcgt tgcgccagcg gcgatcatcc agccgtgtgg gagtagcggg ctggctccga 11940  
 agaattggct caggccggga gtttggacga cagcggccag tgccacgaac gagacaagcg 12000  
 cacacgcat gaccagtggg gtccgtccgc gcacgaccag ggtctgtccg agctgggcgg 12060  
 cgaccagcg cagagcgcg accgtgttg cctggcgag ggttcgggtc gccctgccc 12120  
 ggatccacgc acaaattggc gccccggcgg tgaccgcggc gcggcggtag atgtcgcg 12180  
 tcagcgagc gccagcgag gcttcggggc cttctgccag cagcatctcc gcagtgatcg 12240  
 tcgggggtg tcgcaccgc atcgccatcg ccgggaggat gtcggtgagc aggttgacca 12300  
 gcagcaactg gcgggcgttg aggtgcccc cgcggtgat cacaccggtg acgacggtga 12360  
 agccgatctc gccgaggttg ccgcgagca ggatggccag cgagtcacgc accgatgcc 12420  
 acatggcgcg gccttcaca atggcgtcga cgatggtctc gatgcggtca tcggtgacca 12480  
 cgaggtcggc ggcctcgcg gcggcggggg ttgcgcgctc gccagcgcg atgccgatgt 12540  
 cggccaaccg gatcgccgg gcgtcggttg cgcgctccc ggtcaccgc accacacggt 12600  
 cgttgcggtt caggcgctcg acgatgcgc cctctcgcg cgggctgacc ctgcgaaca 12660  


---

 eegeagcgtt ggtcacttc gccgcagtt ggtcgcgctc gagggtgtcc agttcggtgc 12720  
 cggatcatgac gcgacgctcg ttgagcgcg cgagttcggc ggcgatcgcc tcggcggtgc 12780  
 tgggggtggt gccggtgacc atcacgacct cgacacctgc ctgctggagc tgcgcgacgg 12840  
 cggcgccgc agtgggcgc accgggtcgg ccagcgcgac gaaaccgatc agcccgagat 12900  
 cccggatccg ggactcgtcg agatcactgc ggtccgaggc ggcccgtcc gcgacagcca 12960  
 gcaccggta gccctgcgg gcgaggtttt cgacaactcg ctgcagcct cgacgagcgg 13020  
 tctctccaa cggcacgtcg ccatctgac tcgccaccg ggtgcactgg gccagcacga 13080



tctccggggc acctttcaca ctcagcagat tgccgtcgcg ccaggtgccg agcaccgcgt 13140  
ggtagccacg ggaaggctcg aagggcatct cgtcgatcgg taccagcgc tgcttgccctt 13200  
cggccgggtg cactcccgcg ccacgcgcgc catcgacgac cgcgcgatcg gtgggatgcg 13260  
ggagcagttc tccgtttctc accaacggac ttgcccgcaa ggcaccgcg agcactcgcc 13320  
gttcactgt ggagatttcc tcgagccacc ggctcgacgc gccatcggat acctgccgca 13380  
gcgcgatgcg cccttcgggtg aggggtgccg tcttgctgaa acacaacacc tgcgcccgcc 13440  
ccaaagcctc gatcgtcgcg gaggtagcga ccagcgcccc gcgtttcgac agccgcggg 13500  
cggccgcgag ttccggcgac gtggcaacga acggcaatcc ctcgggaacg gccgccaccg 13560  
ccaggtcac cgcgcgcgc agcgatgtcc ccaatggacg cctacgcagc atgtccgtga 13620  
cgagcagaat cccaccgcgt cccacgcaga tcggcaccgt caccttgcc agcgtctgca 13680  
ggcgagccgc gacaccatcg gcacctcgcc gggcatcggg cccgatcgcg ctgcctgcct 13740  
cgggtgttcg gccggtcgcg accacgacac ccaccgcccc cccggcagcg accgcggtcc 13800  
cccggtagat catgcacgtc cgggtccgca ccggcaccgc cggagtcgcc gtcaccgtct 13860  
tggtgaccaa catcgactcc ccggtcaggc tcgactcacc gacctccagc ccgtcggcct 13920  
ccagaagccg gcagtcggcc ggaacagcat ccccggccac caactccacc acatcgccga 13980  
gcacgagctc gtcagccggc aactgcgtca ctttttccct acggcgacc acaaccgaa 14040  
caacgctggc ctccagaagc ctgttcaaag cgcggtcgc ggtgatgcg tgcacgcgc 14100  
cgatcaacgc gttcatcccc agcacgccc tgatgatcgt cgcgtccgca acggcaccga 14160  
tactcgccga gatcgcgcc ccaccgcca acgcccggc caacgggctc gccagctcct 14220  
ccagcgacgc ccgcaacaca cccatccggc gaaggcccgt cccctcgtcc tgacctgga 14280  
aacgctcgtg cgcgcgatca tcggtcagac cctgcgatga actgccgagc agctccagca 14340  
ccatgccggg gcccatcgca tgccaggcag tgcgcggcac cgcgcgcgga accggacgcc 14400  
acgcgccctg catccccac cagtcccag cgcgcgagc gattgccgca cacacggcca 14460  
ccggaacgcc cgcacggacc tggctcccga tggcaggccc gaacgcgccg aacaccgcc 14520  
cgctgaccga accgatcaac gacagctgcg cgctacgcac gctgacataa cgcgcggtcg 14580  
ggaccgtact caacaacgca cagactccc ccaggtcgc gcccgagatc aggtgcgcac 14640  
cccacggaac ctcaccacca cgcgtcacac caagcccgc atccgcagcc gccagtgccg 14700  
ccgcgcgcgc cgcggacacc acaccgacca cgctgcctc ccgctgcatc tcccgacca 14760  
ccgaagccag ccgcacccca cccgggacca cgcgatcaac accgatccgc cccgccaacc 14820  
ggcccccatc acccgcgaga accaccgacc cgacctcgcg agctgccgag accaactcct 14880  
cagcgagcgg atccagctcc tcggcagcca gaaccacgc ggccgtgcg gaatcgtggc 14940  
gcagcgcgaa cacctcgata ccccgaccgg ccacctcgcg cacgtctggg ggcaaccgcg 15000  
accgagcgac cgagtacagc gaccaaccat cacgtgcg gccgaacccc ttggtcacac 15060  
cacgagttc atgcgcacgt tcccacaact cgacgaccgc gcgcccccg tccgcgggaa 15120  
tcacgtcatg caccaaccga gcaccggtgc gcagaacgtc cgcgtcgatc accaccgtgt 15180  
cgaccocggtc caaccgcgc aacaccggg catcgaacac cagagtcgtc cgctccgaca 15240  
gcccgtaacc gagtgcgcc gcgaacgcgc cccgccccac ctgcgcgcgc cgcggcgcg 15300  
ccgcgaccac catcgctgc gcacgtgaa aactccgct caccgcgagc atgcccgcga 15360  
acgcgcgcgg cccaccaga gccgaaccgt tcgcggcgcg ctcgaccggg ccggcgggca 15420  
acggcacagg acgcggccga acccccacgc gcgcactac gtgactggca ggatgcgcc 15480  
cgttccggc tccccaccgc gccagctcc gcgcgcgcgc ctgcgtctcc cgcaacaggc 15540  
~~agaaecggt aeeegattc gseageaaa tseeggaag egegeaag gtgttseaa~~ 15600  
acgcaactgt aaaccggaac aacaggtcag tggcggtctt cccgatccgc gtttccagcg 15660  
gccccggac ccaggaggcc gactcacgg tcgacaccag cgcgggaaca gacggcgga 15720  
gcacgggcag cggcagcatc cgggccccca ccgcgatcac agcgcgggtc aagttcgac 15780  
ccagcaccgt cgctcagcg atcaccttga ccgattgtc cgggtggctc acaccgaccg 15840  
cagcgaacgg ttctcgtcc aaaccggtgt ccgctccgc gtcctccacc acaccacca 15900  
gtcagcgac cccaaccagg cccgatcgt gccgatgac caccgcccc acgacagcgt 15960

tgatctcgac ccacgdaa cccggcacct tctccaaccg cggggtc tgctcgccag 16020  
 cctcctcagc ccccggtgg tcaacaccgc gcacctcgac gtacgcacgc cctccgacg 16080  
 accacgcccg ccgcccgcgc cccaagcccg cgatgctctg cgcgacagcc accaccggcg 16140  
 ccacgatctg cgaggccaac gcgcccgaag tacgcagcat cgaacacact cagctgtccg 16200  
 acgcaaagcc gaccgacgcc cgtgacaccc ccgcgcgggt cgcggcgca accggccatt 16260  
 cgagccccca accgacccgt accgcgcgaa ccgcgtcctc ccaacgaccc ggcatgcgaa 16320  
 catcgacccg acgaaactgt gccacgacgg gcaggctgac cgcgcgggt cgcaaccgcg 16380  
 ccccttggtg ccgcgcgcct gcgccaggcg ttgtgttccg cgcgtccta gccaccacta 16440  
 cctcccgcgc gcacagatac cgtgagtcac gagcgtagga ccgacgagcg gaaagcgaac 16500  
 ctcgaaacag ccgcacatcc gtccctctcg tggcctgcga gctgatcagc aacaccgacg 16560  
 gcttcgccac gctgttcgac ccgcgcacac accagtagtc ccagcgggtg tgaattcaag 16620  
 gccgcctgtg tgctcgggtt ctccgaaacg gctccgagca tgcagctacg ccggacgcag 16680  
 gcaacgcaac acctacccag gttcacctgg cctgagcgcc gcgcggcgcc gctcagcgac 16740  
 gcctagcccg gaaccgcgta caaacggat cacttcgacg ccggtccgcc ggacgctgc 16800  
 gcgaccgcga caccacgcc aatcgccgcc gccaccggcc agtccatcaa cccaacagg 16860  
 gtcagcgctc caagaccag atagaaggcc atcgtctgat ggtccgggag aacgccccga 16920  
 acggcctcag ccgcgctagc aagatcatca cgagtgggaa ttagtgatc gggacgggtg- 16980  
 aactcgccg ttacgcccgg cagactcacc gtcgttcccc cagcaccgcg cggtttccgc 17040  
 tgcgtagtaa ccacggcacc tcctcaacac agggcccatg ccgcataatc ggctcgcac 17100  
 ggactcaaaa ccgccccaac cgaccaaccg ggcaacacga ttttatctct gaacacgaac 17160  
 gtccccgcgc cgagacgggt ggaatacggg ctgccgccg acctgaacca gtcaggcggt 17220  
 gtgcttcaac aacgttgctg tgcaaccacc tgctcgggtt ggccaaccgt gtggacgcaa 17280  
 cgccggctcc agtgactgct ggtgatcaca ctcggtgatc atgatcgtct ggactgcggt 17340  
 ttgatcaagg caatgttccc gcacctggcc ggggtgcggg agcgggtggt gcgcgggttc 17400  
 gcggagttca ccaacgagta tccgtggcgg tggacaccgt cgcgtgtgga cgaatggtcg 17460  
 caatcaccga ccggagaacg gcctctcgcg ccatccacaa tccgcgccta ccaaggcagt 17520  
 ctgaggctgt tcaacgaatt cctctgtgac agccggatat gctgggcgct ggctgagcag 17580  
 gcgtcgttcg ggccggacga gtaccgggtg gcgaatgtgc acgagtggaa ctactcgcc 17640  
 catctccagt cctatgaggg cgatccggac gaccgcccgt tcaccgcga tgagctgcag 17700  
 cggttccttg actacgccac gaccaagtgc accgcgcggt tcgcgcgaaa cgcaaccgcg 17760  
 cattggcagc ctatcgagac gcgacctgt tcaaagtgat gtatcgctag ggcttggtcg 17820  
 ccggggccct acccgcggtg gctcccgca ttttttgagt ttgtgcagtt cagaggacga 17880  
 tgtgggagat gctagccgtc gcgacagatg tgggtccagt aggcgcggca ggcaaacctc 17940  
 agcggctcgc gtgcagtatg cggctcgcgc gcagtgcgcg ggggtggttcg cctcgaaac 18000  
 ctgcggatcc gtggaggtag ccggcgatga ggtccggggc ggagtgggtc togatgcgg 18060  
 tgaaccgag ggctgaagc tgtgcggcga tgcgtcgggt tgtgaagtag ctgaacaagg 18120  
 gttcgccggc gccggccagc cgatcggcac gtgcccgtg gtgcgcacgg tctcgtcgg 18180  
 tggccgcagg ctgcaggtag tcgaagacca cctcgaccgg ctgggcttga ccagcgatgt 18240  
 attcgagggt ggctggggc gcgttcgcgg tcagatagaa gacaacgcc agccacacga 18300  
 atactgccg gtccgtccgt ttgaatccg cggattccaa tccagttgcc agcggctggg 18360  
 tttcgaaac gaccggtacg aaggtaacg tctcaggcg gtcgatccct gcgctagcaa 18420  
~~ggcgttcgcg tttccatgct tgggtggcgg ggtggtcgac ctggaagacg cgaagtccg 18480~~  
 ggccgcggatt gcgataggcg aagggttcca ggcccgccg gaggatcacg acctgccgta 18540  
 caccatcggc gatggccgc gccacggcgt cctcggcgaa gcgggcacgg gcggcgaaga 18600  
 acaggcggcg gggctgatcg attgccccg gcgcgaggtg atcggttggt ggctggcca 18660  
 atttggtcag ttcctctgc gtgacgcga gcagacgtac cgcaggggg tcggtgaaga 18720  
 tctgtggctg gtcagcgatc tgggtgatagg cgcgggcgta cgcggtagcg agcgtgtcc 18780  
 ggctgggtcc gccgttctcc atgttcaaga acttactcac gaacgcggcg cgggctctat 18840

gaactcactc	acgcccgcctt	cacaaaggag	cacgcgaaca	cgatccgtaa	ccacagctgg	18900
ccctccttta	ggcgattgc	cacgagtcga	caggcagggc	tcgaaccctg	ttatcggggc	18960
tgggttgatg	tggtgagccg	ggatcatgagt	tcgcggttgg	tggccccggc	gcggcaagat	19020
cgccggcgcg	ttcgctcgagt	tgctgttgct	gcggcaagaa	acgacagctc	ggccgaccgt	19080
tcaacgggtca	cggtgcgttc	ggaccgcgtc	ctgccaccac	acctgacggg	ctttgtcgtc	19140
accggggccct	ccccattgat	gtggccgctg	atgctcgaat	gcgcacgagt	ggtcagtcaa	19200
cgccggttcc	cgagcccgc	cttccccccg	gaagacggcg	cggtcccgca	cggcagcccc	19260
acgcgttcac	ggggacgatc	atcggtccac	gaggcgaata	cctgctccgt	gcaggtggaa	19320
gcggctattc	cagggggccg	gccgtggttg	ctcgaatcgg	cgctcgacg	cggctgacgg	19380
tgggcgcggc	atgatgagcg	accgtgctgc	tgcgactggc	cttgccatg	ccgataacct	19440
cgcccggtga	ggtcacacag	cagcgtctga	cggttttccg	gtcgagcccg	gccagccggg	19500
tcacctcag	cagcgaatgg	cgcgccagcc	acagctggct	tggttctcag	cggttgaaca	19560
ccggtagctc	ggctggaggt	tgccggatga	atccgccgtg	ctggggccatt	tggtcgtaga	19620
gatgaccgtg	cgactccgcg	agcgttgcaa	gcgctccgtt	gacgtcgacg	tgaccagacc	19680
ggccggttcc	cttgaccacc	cgcccatcga	ccagcaccgt	gtgcacgttg	cccaggttgg	19740
tctgcgcaat	gatcgcgcca	caagggtcgt	ggcggttcca	accggcctgg	ctgatgcctg	19800
acatgtccag	caatacgatg	tcggcccgcgt	tgcccgggtgc	cagggttcct	gtcacgtcgt	-19860
cgacgcccgc	ggcgcgggca	ccgttgaccg	tgagccaccg	taaggcgctcg	cgcgttttcc	19920
accgcatgac	cgtgggcagt	gtcgaccgcg	cgtagctctc	ttgatcgctc	ctccatcggg	19980
tcacctgtag	gacgagccgt	gcgtgcgaaa	tcattgtcgcc	gcaggtactg	ctcgtgcagt	20040
cgatgccgag	gcttggtccc	ggcgtgttct	cggtggcctc	acgaatgacc	gggaatccca	20100
tgcccatctg	catctcggtc	tcggcgcaaa	ccgagatcat	ggtgccggta	ccccgcagca	20160
ggtgccattc	gttctcagtg	ctgaatgtcc	cgtgcaccaa	gagaaggctcg	ctgccgagca	20220
tggttggtgga	gtgcagcgcc	tcgatgtcct	tgaacaggtt	gcgcgcgatg	acttggttgc	20280
agtggaaagt	gatgcgtgcg	ccgagctcgc	gggccagctc	gaattcctgc	ttgacgtccc	20340
cgaagggagc	gatggctagt	tcctgcgggg	cgatgccgaa	gcgcagcaat	tggtgctcgc	20400
tcgggaagta	ctgggcgcgg	atttcccggg	ccagccgcgc	ccgccggggc	aagtgcctt	20460
cttccaccga	ctcgtccagc	ccgccttttg	tctcggacca	cgtgttcgac	gtgatgggca	20520
ggaggccgtg	gccgtacagc	gcgcgcacgc	ccgcgtcgcg	gagtcacagc	accgcggcgt	20580
gcgcgtgctc	gccggtgacg	atgttggtgg	agtaatcgac	cagcgtggtc	acccccgagt	20640
taaggcaatc	gaggccaccg	aggtagttgc	ccgcatacat	gtcctgcggg	cggtaacttcg	20700
tcgcatctg	cagccggaat	ccctcaggt	agtcgaggat	gttcccgtcc	gccagaatgc	20760
cccggagccc	ggtttgccag	gtgtgtcgat	gggtgtcgac	catgcccggc	atggcgatca	20820
tcgacgaggc	gtcgattcgc	tcggtgtcgc	caacgttgag	gttcggcccc	atctcgacga	20880
tgacaccgtt	gtcgacgagg	atgtcggctc	gttcgaagtc	gccgagaacg	tcgtccatcg	20940
tcaccacggc	agcgttctcg	atcagtgtgc	gcattggccaa	gtgctcctct	ccgcgcgggtg	21000
ggcatcggat	gccgtcctca	agcgggtccc	ttcgggtcgtt	gatcgactgt	ggagaagagg	21060
ctgtggagag	gggcccgtct	gcagacgggg	cagcttctac	tggtccacgc	caatccgaaa	21120
accggcggtt	gtcagtcgca	tcgtcgcgcg	gtcttctgtt	gaatgaccca	ggtcgcgatt	21180
tgctgccgag	tggtgaagcc	gagtttggtc	aggatgtgct	cgacgtgggc	ctcagcgggtg	21240
cggcgggaga	tcacctgat	ggcggcgatt	tctttgttgg	tggtgccctg	tcgaccagc	21300
gctgctacct	ctcgtcgcgc	cggagtaagc	accgacgcgt	gcgtttcttc	ctcctggag	21360
cttgctggtt	ccccggcttt	gcccggattc	cgtatccatg	cgatggcctc	gtcgaagtgc	21420
agctgtgtgg	cgtggtgaac	gcctttctga	aacgtctcgt	cgccgagtga	gctacgtgct	21480
actgctcgt	agtggctcgt	taggcgcgca	acatgtttgt	ggccggcgag	cgacgcacgc	21540
acgttctttt	tgatcgtttg	ggacacgccg	aacaaggctg	cagcccgtgc	cgggtcaccg	21600
tctgcgctgg	ctatgcaggc	cagaatctca	acgcacaacg	cgatctgaaa	ctgctcatgg	21660
aagccgagcc	gcaggcgaag	actctccagt	tccgcagcat	tggttgccg	gtaattgcc	21720

gctagtaggt ctgctattc cttcgccac agggcaagcg atctcacc ca ctgttcgctg 21780  
taggtctcct tgacggcagg gcaccgcgg aacagggcag cggcgcgctc atgatcacca 21840  
aggaacgcga ctgccaacgc gagctcgatc tgatcgtagg cggcggcagc gtgatcgcca 21900  
atcagccgat gcttggcgag agcgtcttcg aagagcgcgc tcgcccgcgc cggatcttcg 21960  
ttgaacaggg cggtagagccc cgagatctgc ggcacgtagg ccgcccctga ttggtcgccc 22020  
aactccgccc ccagcgcgct cgcctgctcg agaagggtg acccgatga cgcgtcggtg 22080  
agcgcggcat tcaggtaccc gtcgacgtag agagctttga cacgtaccgg atccgcacgc 22140  
aatcgatggg aacoggacat atcgggctaa tgggtgcacaa cgggcagtgt tgatctgacc 22200  
acgctaaccg gtatctcggg cgggtcgctg gctgcgccgt tcaagctcaa gcgtgcggtg 22260  
cgtcttcgcg tcgacgatct acaagtacgt ccccgaaactc agcaccacc gagcgtcgc 22320  
ctctgggccc ggtccgcagc cagcgttcca gagctaagcg ctggtttccg gcgtctgcta 22380  
ccgggacgct tcggccgccc tgggttcagc gcaagtgagg gttcctcggc gggaacccaa 22440  
gggtccgtcc cggccaggac caggactgga cccgtgggtt cggctcaacc cggagtgtcg 22500  
catttgatcg tgggtttccag cagtaatgga tgctgaatct tgtgattgac gccgccgcga 22560  
ccggagagtg tcttgagca accgattccc gcgcgattcc ggaccatcca tcggacctag 22620  
gttctgttgc gacctgagcc cgcggcgcg gcggagtttc cgcctcaagt ggccaggctc 22680  
agccattcct cgatgggcag cacgtgtgcg cgggtgtccg cgtgcgtgcg gaactcattc 22740  
tcggggtcgg agacctcgac gtagcccgcg atcttgcgcg gctgttcgcg ctcgtagtgc 22800  
atcggcaactg cgtaccgggc gtcgaggatc tctgcggctg cggcggcctg cctcgggtcc 22860  
atcgcggcgg gcagcgggct cggcggctgg aggtgtggcg cgtcgaccac cgcgccgttg 22920  
gcgggcagga acaccgcgct gaacgggctg aaccggcgcg cgatgagcca ccagtagccg 22980  
tggaacatcg tgtcgccgccc gtggaagacc cgctgcccg cggcctgcac caccagttc 23040  
agctgcgggt cggccagccc gtcgacggcg gggaccgcgg tgacgcggaa cggcccagg 23100  
tcgcgggtgg accaggcatc cagcacctcg gcagccagtc ggtgcagtgt cagctcgcgc 23160  
tcggccggca gtgtcgtcac gttgtccacg tcgtcgccgt ggccgggcgc gggtcggagc 23220  
accggtgccc ccggtgtcag cgcggccgcg agcgcggccg cgtcgggtgtg gtcccgggtg 23280  
aggtgggtga ccagtgcggc ggtgaccgtc ccgctcggca ccgccagcct ctcgccgggt 23340  
ttccacgccg tgaacagcgg tgagaggtcc cgtacgtagt cgatcaccag tcgcttgccg 23400  
ccgcctcga tttccagtcc ggcccagccc agtcgtcgca cccgatctc gcactccttc 23460  
ggtccgtggg tgaagcggaa tttagcgtac gcacgttcac taaatcaata gcgaacgcgc 23520  
gtacgctatc ctggtcacat gtcaccgcga cgctcagcgg ccgaagcaca ggccaccagg 23580  
agccggatcc tcgaccgtgc cgcgcagatc gcatccgagg aagggtgga cggcatcacc 23640  
ctcggccggc tcgcccagga actggagatg agcaagtccg ggtgcacaa gcacttcggc 23700  
ccaaggaga cactgcagat ctccacgctg gacaaggcat tcgtggactt ctggcaccgg 23760  
gtggtcgagc ccgcaactgg cgagccgccc ggtctgcggc ggctgcgcgc ggtgtgcgcc 23820  
aacgccgtgg gctacctaga agcgcactg ctgcccggcg gctgcctgat gaccgcggcg 23880  
ctcaccgagt acgacggccc ccccgccggc gtcgcgcagc cgggtggcca ggtgtggtcg 23940  
cgctggcggg gacagctgcg ggcggacctg accgcggcg tggagaacgg cgagctgccc 24000  
gccgggttcg acatcgacca ggcgctgttc gagatcgtcg ccaccgggct ggcaactgaac 24060  
gcggccatgc agctccagca cgaccgggcg gccgcggcc gggcccgcgc cgcgatcgaa 24120  
cgggccttgg accagtctg acgcccgcgc cagacaaacc cggccgattc tttcccagg 24180  
~~aseagacgt ctgaecggt teetgetggt eageecgtat ceateceatt tttcggegag 24240~~  
tacgccctag catggtcacg agacattccc ggagccgtca ttgattgact ggcgcagcga 24300  
aaagctactc gccggttgct ttgcacctg ggtcagcttc gggagtccg ggcccagagt 24360  
ccggtgggtt cagtgtgagt gaaggttcct cgctgggagc ccgaggggga ctgccaccag 24420  
ggtctgaaga ctctgtgaag aaggtatgaa cccgggttgg gcatgcttcg ttatgagggt 24480  
atgggaagcc aaggcttggg agcgtggaca tcatggcaaa cgatcacggc acggtgctcg 24540  
ccaccatgat cttcaacctg gtagccccgg cgggtgtgat cgcgccagtc ggcgtggagc 24600

tgcggtacga	cagccgcaat	ccgtacgaga	tctccatgaa	gctcaacgta	ggcacggacg	24660
gtcaggtgga	ctgggtgatc	gcccgcgacc	tgctggccga	cgggctgata	gccgaggcag	24720
gcgaaggcga	tgtgcggatc	ggccctcgac	ggggttttcc	ggggttggtc	gtgatcgaga	24780
tgagctcgcc	gtcggggcag	gcctccttcg	aggtgaatgc	tgaccagctt	gcggacttct	24840
tgaacgacac	ctacgacgtg	gtcgaacctg	gtgatgaaca	ccggtggatg	aacgtcgacg	24900
aggtgctgag	ccagctgctc	tcgcacaacc	tgtaatggcc	cagctctccc	gaagcgccgc	24960
acgccaaagc	gctggctgcg	ggacctggcg	gcgctgaaca	ccgccacgct	gtgtctccga	25020
gctccagctg	gaccacgtcg	gtgccgtgcg	cccggtcgcg	tcaggccgaa	ggtgctgata	25080
ttctccaggc	gcgccatcgg	cgcaggaagc	gctgcttctg	ctcccgcgcg	agtaccgtcg	25140
tgtcatggcc	acggacagct	tcgattcctc	gaagctacag	gcggccgtgg	catcgagcgt	25200
cgcgtcgtgc	gtctcggaag	tcagccgaga	cgtctacacg	cacctgatta	ccgaggctcc	25260
gcagttgca	gccgatgaga	tcgtcctcag	cattctacgg	acgagtgttg	agggaaatat	25320
cgccacattg	ccgcacgttc	tcgaattcga	gattccgctg	ggatattcgc	cgggtcctgc	25380
tgcggtgttg	gagtatccgc	gacgactggc	gaaacatttc	catcaacgcg	ctgatcaggg	25440
ccaaccgcat	cgggcacttc	cgcttctgt	agtgatgcct	cgacgagatc	cgcgcgcaat	25500
gcgccgacga	ggccgtatcc	gcagcgacca	cgcaacgaat	gctcgcaacc	agcttcggct	25560
acatcgaccg	cgtcacggag	cagatcgccg	aaacctacca	gctcgaacgg	gaccgctggc	25620
tcttggcgac	gggacggccg	tgaggtctct	gcggcatccg	catagcgtct	tctcccgtct	25680
aggcacatga	ggtgtttgcg	gcggtcgttt	ccggcagtcg	cacggcattc	gtcctagctg	25740
cgggcaattg	agggagcgaa	gatttagagg	agtgtggcca	cgcggacca	gccggcgagt	25800
gctcgggagc	ggctgtgggg	cggccaggcg	atgactgtcg	tcacgtccgg	cgcgtctaga	25860
accggtacgg	cggcgaggcc	ttcgagcagg	ttgacgcgac	tggattcggg	catgaccacg	25920
gtagtgcggc	cgagtgcgat	catttggaac	agttgctgt	ggttgcgtac	ttccacgccg	25980
gggccatctg	gatagacgcc	gtcggggccg	ggccagcgcg	caagcgggag	atccggcagt	26040
gagctgacat	ccgccatccg	tacatggggc	tcgctggcaa	gcggatgcga	ggtcgggaaga	26100
atggcgactt	gttgcctcgt	gttcagaatt	tcgatgtcga	gttcggccgt	cgggtcgaag	26160
ggttgatgca	acagcgccac	gtcggcccgg	ccgtcatgca	gcgttttctg	gggctgggat	26220
tcgcagagca	gcaggtcgac	ggccacggct	cccggtcgcg	cggcgtacgc	gtcgagcaac	26280
ttcgccagca	gctcaccgga	ggcgccggcc	ttggcagcca	ggactagcga	gggctggctc	26340
gtcgcggcac	gctgggtgcg	tcgctcggct	gctgccagcg	cgccgaggat	cgcggcgct	26400
tcggtcagca	gcattgcccc	ggcttcgggt	agcgagactt	tgcggtggt	gcgttgacgc	26460
aacacgactc	cgagtgcgtt	ctcgagctgg	gcgatcgctc	gcgacagcgg	cggctgggcg	26520
atgcccaggc	gctgggcggc	ccggccgaag	tgcaactcct	cggcgactgc	aacgaagtac	26580
tgcaactccc	gcgtctccat	ccgtcgagcc	taccgctgat	tcatatcagc	tgggtatcgg	26640
tgtgagacct	agatggtgtt	ggttccccgc	cggtttcggg	ccacgctaga	aagcatgagc	26700
gaacagacga	ttgcaactgt	caccggcgca	aacaaggga	tcggatacga	gatcgcggcc	26760
gggctcggcg	cgtcggggtg	gagcgtcgga	atcggggcac	gggaccacca	gcgcggggag	26820
gatgccgtgg	cgaaattgcg	tcgggacggc	gtcgatgcgt	tcgcggtatc	cctggacgtg	26880
acagacgacg	cgagcgtcgc	ggctgctgcg	gctctgctcg	aggagcgcg	cggccggctc	26940
gatgtgctgg	ttaataacgc	cggcatcgcc	ggggcatggc	cggaggagcc	ctcgaccgtc	27000
acaccggcga	gcctccgggc	ggtggtggag	accaacgtga	tcggcgctcg	tcgggttacc	27060
<del>aaegetatgc</del>	<del>tgeegttget</del>	<del>aegeegctec</del>	<del>gagcgceegc</del>	<del>ggatcgctca</del>	<del>ccagtcacgc</del>	<del>27120</del>
cacgtcgctt	cctgacctt	gcaaaccacg	ccgggcgtcg	acctcgggcg	gatcagcgga	27180
gcctactcac	cgtcgaagac	gttctcaac	gcgatacca	tccagtacgc	caaggaactc	27240
agcgatacca	acatcaaaat	caacaacgcc	tgccccggct	acgtcgcgac	cgaccttaac	27300
ggcttccacg	gaaccagcac	gccggcagac	ggtgccagga	tcgccattcg	gctcgccacg	27360
ctgccagacg	acggccccgac	cggaggcatg	ttcgacgacg	ccgggaatgt	gccctggtga	27420
ggcgctcagt	cggcgatggt	gcaatcgaag	tcggagaggc	tcgctgcgac	cgggtacgcc	27480

gaacaacacc tgttctctg ggtacggatg tcggccttcg ccgtctctgt cattgacaac 27540  
 ctgtacttcg ggcgcggtta ccgcgcgtgc gccgcggttg cctggcgaca ctgggccagc 27600  
 cgtgggtcac cggcgggtta ggtcaggcgt gggcgggttg cagcatggcg ggtgcggtt 27660  
 tgcgtaggtc gggtaggcgc atccggcgcg ggagccggtc gagttcttcg ccgatggccg 27720  
 gtgctttggg gctgctcagg agccgaacac ctcccagccg caggtgccgg gctgaaccga 27780  
 gtggttctcg tcggctcggg tcacaacgtc tgccggaaca gctgcggcga ggtggctcga 27840  
 gattcgaggc gggatcgctc tcggcgacct tgccgacgat cgcggctagg gccagggtt 27900  
 tcgtcgacct ggttggcacc tagatcacga cggtaaaaac ttgccggcat cagagacgat 27960  
 cgaagtgatc ccgggtcacg tcggcttacc ggtcgagtga gtcccggggc ctgccagcc 28020  
 aggtcttgcg tcgttggttc gggctcagtt gcggattccg acgaacaggc ctcgccgtt 28080  
 cgggtgctcca ggaaggattt ccgcgcggat ccctgcgtct tcgagcgcgg cgggtgtactc 28140  
 gtctcagtg aacagcgaga ggatttcgaa ctctgtgaag tcccggatcc cggtggttc 28200  
 ggcgactgtg tagcgacgg tcacccggct cgtacggccc tccaggaccg agtgcgatag 28260  
 ccggctgatc acccgctcgc cgtggtgccc gacggctccg gtgacgaacc cgtcgatgaa 28320  
 cttgtcggga aaccaccagg gttcgatgac cgcgactcca ccaggggcca ggtgccggc 28380  
 catgttcgc gtcacgcgtc gcaggctcgc aacggtccgc atgtaagccg cggtaaagca 28440  
 caggcagggtg atgacgtcga atggctcgc gaggtcgaaa tcgcgatgt caccgatgtg 28500  
 aatcggtaac tcagggactc gtctgatcgc gatctccgc atcgcatcg acagttcaag 28560  
 cccgcgacc ttcgcgtatt cggcacggaa tcgctctagg tgcgccccg tcccacaggc 28620  
 gacgtcgagt agggactgtg cttcgggcag cctggtgcgt acgagctgga ctacttccc 28680  
 ggcctcggct gccagtcgc gccacgcgc ggagtggatc gcgtcgtaga tgcggcatg 28740  
 atctgggctg tataccgagg aggtttctgc gaatgtgctg ctacgcgcg acatcctcac 28800  
 tttcggagtg gtgatctttg gctgatgtgg tgttcgacgg ccttctggaa ctgctcagcc 28860  
 accgtgcga cctcggcgtc gtcaaggctt ggtgacgtg gtagcaggag tgttctgcg 28920  
 caggcgtgct ccgcagaagg cagcttgacg tccgcgcggt agatggggac cttgtgcagg 28980  
 ggcgggtagc ggtagctcgt gtagatgccg cgttccagca tttgctgcgc cacctggctg 29040  
 cggatctccg gagccagctg gaccagtag aagtagtgag acgagacgtg cccatccggt 29100  
 agcgtcggcg gtaggaggac acccggcaca tcggaagca accggtcgta ctgcgtagcg 29160  
 atttctctac gcctgttgat gaattctggc agtttgcgca gctgcacgct gccaaagcgt 29220  
 gccgtcatgt cgttcccgat cagccgctgg ccgatgtctt cgacgcgaat atcccaccag 29280  
 cggttggaag acttgccga atcgaatccg ctcatctgct caagaccgtg gtaggcgagt 29340  
 cgtcttgccg ggtgcgccag ctccggatcc gccgcgtaga acatgcccc atcccgggtg 29400  
 accaggatct tcacgcacg gaaactccac gtggccaggc caccaaaggc tccgaagcg 29460  
 gtgcccgtgca cggacgatgc caccgcgcag gcggagtcct cgatgagcat gaggccctt 29520  
 tcacggcaga aatcggcgat cgcggtgact tctccggcg atcctccata gtggagcagc 29580  
 aatacggcct tggtcgccg cgtgatggcc ctgccacat catccagcgt ggggttcaac 29640  
 gtccgggggt cgacgtcgca gaacaccggg cgggcaccgg aggatgcgat ggcgttggc 29700  
 gccgccacga agcttatcga aggaagtacc acgtcgtcgc ctgggccgag gtcgagcacc 29760  
 tgcacggtaa ggaacagcgc ggcagtcgcc gagttgagga acacgacctg ttcgggatcc 29820  
 actcccagg ggtgggcgaa ttcggcctcg aacgtccggg tgcgcggccc gagcccgatc 29880  
 cagttggagg cgaacacctc ccgatcgcg tcgagttctt cggtgccgag gatcggttg 29940  
~~tgcaggttga tcaegttgct gaaateetee gagatgeege eatgetggat getaggaaet 30000~~  
 cttggccacg aattcagcga ttgattcgac gacgtagtcg atcatttggt ccgttatgcc 30060  
 tgggtagacg ccgaccaga aggttcggtc ggtgacgatg tcgctgttgg tgagcgcgtc 30120  
 ggcgatccg taccgcacct gctcgaaggc cgggtgccgg gtgatgttac gccgaacag 30180  
 cagtcgggtg ccgatgttg gggattccag gaagttcacc agggcggcac ggggaaccc 30240  
 ggcgtccgca ctgatgtga tcgcaaacc gaaccagctc ggtcgtgtg gcggtgtggc 30300  
 taccggcagc agcaggccc gcaaccggga cagcccttcg cgaaecgtc gccagttacg 30360

gcggcggtgcc	gacccgaatg	cggaatatctt	gctcaactgg	ctcagcgcaa	gtgcggcctg	30420
caggtcggtg	gtcttgaggt	tgtaacccgac	gtgggagaac	gtgtacttgt	ggtcgtagcc	30480
cgggtggaagg	gtaccgaggt	ggtagtcgaa	cctcttgccg	caggtgttgt	ccacgccggg	30540
ctcgcaccag	caatcccgtc	cccagtcacg	cagcgactcg	atgatgcgag	ccaattccag	30600
gctgccggtc	aacacgcagc	caccctcgcc	gctggtgatg	tgatgggcag	gatagaagct	30660
gaccgttgtc	aggtcgccga	aggttccggt	cagccgtccc	cggtaggtgg	atcccaccgc	30720
atcacagttg	tcttcgacga	ggaacagctc	gtgttctttt	gcgatctccg	cgatttcgtc	30780
agcggcgaaag	gggttgccca	gggtgtgcgc	cagcatgatg	gctcgcgtcc	gttccgtgac	30840
ggcggcccttg	atgcggtctg	gcgttgcggt	gtaggtgccc	agttccacgt	cgacgaatac	30900
cgggacgagt	ccgttttgga	ccgcgcgatt	gatcgctcgtg	gggaagccga	ccgccgcagt	30960
gatcaattcg	tgcgccggcc	gcagtcgtgc	ctcgccgagt	ttgggggagg	taagcgaaact	31020
cagtgccagg	agattggccg	acgaaccgga	gttgacgaga	tgagccttgc	ggaggccgaa	31080
gaagcggggcg	aactcgctct	cgaatcgccg	tgcattcccg	cccgcggcga	tccggagctc	31140
cagcgcgggt	tccaccagtg	ccaccgcgtc	gtcctcgctg	agcacggcgc	ccgatggccg	31200
gatcggcgtc	gatccagcca	cgaaggtcgg	ggattcctgt	tcgcggtggt	aatcgcgtac	31260
ggatgccaat	atccggtcct	tggcatccgg	caccatctca	gtagcggtag	cgcaagtgtc	31320
gtcacacgaa	gtcactctgg	cgcgcctttt	cccagcgct	ctggttttcc	ggctctgcat	31380
gcagggcgacg	atcagctctc	gcgccttgcc	ttcaggagat	gagcgatgcc	cgtggcgaaat	31440
cgcgttatga	cgtcccagcg	ggacagtgtg	ctgtctcggc	gccttacacc	ttcctgccct	31500
ggttcgatgc	ggtgcgggac	atcaggacag	cggagcaagg	agaagcgctc	attgactcag	31560
aaatcctcga	tctaccggc	acaccgcact	cggtagagcc	caggctagcg	ggaacgacct	31620
gctcgcgctt	gtcaagatcg	ctaccatcac	ctggaaggcc	taagatttgg	cttgcgaaag	31680
cggcggtttcc	cgggggatat	cagagatttc	tgtgattctt	ggcatgcttc	ccgggtgttc	31740
aattgcgata	ggagagttca	tgcgtgtcct	gttcaccccg	ctgccggcga	gttcgcactt	31800
cttcaacctg	gtgccgttgg	cgtgggcgtt	gcgtgccgcg	gggcacgagg	tccgtgtcgc	31860
catctgcccg	aatatggtgt	cgatggtcac	cggagcagga	ctcaccgcgg	ttcccgtcgg	31920
cgacgagctc	gacctcatct	ccttggcggc	caagaacgaa	ctcgttctcg	gcagcgggggt	31980
ctcgttcgac	gagaaggggc	ggcatccgga	actcttcgac	gagctgctgt	caatcaactc	32040
cggcagagac	acggacgccg	tggagcaact	ccacctgtg	gatgaccgat	cgctggacga	32100
tctcatgggg	ttcgccgaga	aatggcagcc	tgatctcgtt	gtgtgggacg	ctatggtgtg	32160
ttcggggcca	gttgtggcgc	gagcgctcgg	cgcacgacac	gtgcggatgc	tcgtcgccct	32220
cgatgtgtcg	gggtggctgc	ggtccggttt	cctcgaatac	caggaatcga	agccgcctga	32280
gcagcgcgtc	gacccgctcg	ggacgtggct	gggagcgaag	ctcgccaagt	tccgagccac	32340
ttcgatgaa	gagatcgtga	cgggccaaag	gaccatagat	ccgattccat	cctggatgcg	32400
cctgcctgtg	gacttggact	acatctcgat	gcgtttcgtg	ccgtacaacg	gtccggcggt	32460
gttgccggag	tggttgcgcg	aacgaccgac	gaagccgcgc	gtctgcatca	cgcgcgggct	32520
gaccaagcgg	cggctgagca	gggtgaccga	acagtacggg	gagcaaagtg	accaggaaca	32580
agcaatggtg	gaaaggttgt	tgcgcggcgc	ggccaggctc	gacgtcgagg	tgatcgccac	32640
cttgtctgac	gacgaagtac	gggagatggg	ggagttgccc	tcgaacgtcc	gggtccacga	32700
atacgtaccg	ctcaacgaac	tgctggagtc	gtgttcagtg	atcatccatc	atggctcgac	32760
gacgacgcag	gaaaccgcca	cggtaaacgg	cgtaccgcag	ttgattctcc	ctgggaacct	32820
<del>etgggaegaa</del>	<del>tetegttagg</del>	<del>eggagctset</del>	<del>agegatcgg</del>	<del>ggagcgggtc</del>	<del>tggctctcga</del>	<del>32880</del>
ccccgcgacg	tttaccgaag	acgacgtgcg	aggtcagctg	gcccgcctgc	tgcacgagcc	32940
gtcgttcgct	gccaacgcgg	cgtgatccg	ccgtgaaatc	gaggaaagtc	ccagcccgcg	33000
cgacatcggt	ccacgtctgg	aaaagctagt	tgccgaacgt	gagaaccgcc	gcaactgggca	33060
gtctgatggc	catccgtgag	caacgtgtgg	ccggaaacat	ggacgccggg	gtttggcagg	33120
tgttcatcgc	tgttgcgctg	actcggatcc	cgcgcgtgacc	gggacgatgc	caggcgagtc	33180
ccgaagtcag	attcttgtcc	agaatcgtcc	aatgggggtgt	tgatctcccc	agaggtttgc	33240

gctccaaccg atttccg aggatcgtgg cgcccgtga gcaaccg cagtgcggtc 33300  
gagacatacc gctgtgcgcc aggagcgaag gtgggttgcc cgatcacctg gctgggtgta 33360  
gatgccgagc cgaaggtcac cttggatgag gcggaagcct ggagagagca caccgaggcc 33420  
gtggccgacg tccgtgtctt ctccggcggg catttcttca tgaccgaacg ccaggacgag 33480  
gtgctcgccg tccctacggg cggatcgctt cgatgatcct cgccaggccg ctggaccaga 33540  
ccgcgacgcc cctgggagcc ggcgtgcaca tcgtcacggc agtgagggat tgggcatgag 33600  
cagttctgtc gaagctgagg caagtgtgc tgcgccgtc ggcagcaaca acacgcggcg 33660  
gttcgtcgac tctgcgtga gcgcttgcaa tggcatgatt ccgaccacgg agttccactg 33720  
ctggctcgcc gatcggtgg gcgagaacag cttcgagacc aatcgcatcc cgttcgaccg 33780  
cctgtcgaaa tggaaattcg atgccagcac ggagaacctg gttcatgccg acggtagggt 33840  
cttcacggta gaaggcctgc aggtcgagac caactatggc gcggcaccga gctggacca 33900  
gccgatcatc aaccaggctg aagtaggtat cctcggcatt ctctcaagg agatcgacgg 33960  
cgtgtgtcac tgctcatgt cagcaaagat ggaaccgggc aacgtcaacg tctgcagct 34020  
ctcgccgacg gttcaggcaa ctcgagcaa ctacacgcag gcacaccgtg gcagcgttcc 34080  
gccctatgtg gactacttcc tcgggcgggg ccgcggccgc gtgctggtag acgtgctcca 34140  
gtctgaacag ggtcctggt tctaccgaa gcgcaaccg aacatggtg tggagtcca 34200  
ggaggaagtg ccagtcctgc cagacttctg ctggttgacg ctcgccagg tctggtctt 34260  
ccttcgtcag gacaacatcg tcaacatgga caccggacg gtgctgtctt gcatcccggt 34320  
ccacgattcc gccaccggac ccgaactagc cgcctcggag gagcccttcc gacaggcgg 34380  
ggccaggctg ctctgcacg gcatcgattc gtcgagtatc tccgaggcgg tcggttggt 34440  
cgaggaagcc aaggcccgct accgcttgcg ggcaacgcgc gttccgtga gcagggtcga 34500  
caagtgtat cgcaccgata ccgagatcgc ccaccaggac ggcaagtact tcgcggtgat 34560  
cgcggtgtcg gtgtccgga ccaatcgtga ggtcgccagc tggacgcagc cgatgatcga 34620  
accgcgaaa caaggtgaga tcgcactgtt ggtcaagcgg atcgccggag tgctgcacgg 34680  
tttggtccac gctcgggtgg aggtgggtta taagtggact gcggaaatcg ctcaccgg 34740  
ccagtgcagt gtggcgaact accaaagcac ccgctcgaa gactggccgc cgttcttga 34800  
cgacgtgtc accgccgat ccgaaaccgt gcggtacgaa tcgatcctgt ccgaagaagg 34860  
cggctcggtt taccaggcgc agaacaggta ccggtatcag gaggtgcatg aggacttcgc 34920  
ggcacgacct ccagcgact tccggtggat gactttggga cagttggcg agctgctccg 34980  
gagcaccac ttcttgaaca tccaggcgcg cagcttggtc gcctccctgc atagcttg 35040  
ggcgttggg cgatgaccag ctcgatgcga aagccggtgc gcatcggtgt gctcgggtgc 35100  
gcttccctcg cgtggcgacg gatgctgccc gcgatgtgc acgtggccga aacagagggt 35160  
gtggcggtg cgagccgtga tccggcga gccaacgggt tcgcagcgcg attcgaatgc 35220  
aggcggtgc tgggttacca gcggtcctg gagcgccgg acatcgatgc cgtctacgtg 35280  
ccgttgccgc ctggcatgca tgcagagtgg atcggaagg cgttgaggc agacaaacac 35340  
gtgcttgccg agaaaccgt gacgacgac gcgtccgaca ccgctcgctt ggtcgggctg 35400  
gccaggagga agaacctgt gctgcgggag aattacctgt tctccacca cggccggcac 35460  
gacgtggtcc gcgacctgt gcaatccggg gagatcggtg agctccggga gttaccgcc 35520  
gtgttcgaa ttccgccgt tccgacacg gacatccgt atcgaccga actcgggtgc 35580  
ggagcgttgc tggacatcg tgtctatccc gcccggtgcg ctcggaact tctcctcgt 35640  
ccgctcacg ttctcggcgc aagctcgac gaggccagg agtcgggct cgacttgctg 35700  
~~ggcagcgtg tgetcaate ggaaggtg gacgttgcg aactcgata ggttctggtg 35760~~  
caccactacc gcagcgcgta cgagctgtgg gggagtctg ggcgaatcgt cgtcgaccgg 35820  
gcgttcacgc cgcgcccgga gtggcaggcc gtgatccgaa tcgagcgga gggcgttgc 35880  
gacgagttgt ccttgccagc ggaagatcag gttcgcaagg cggtcaccgc cttcgacgc 35940  
gacatcagag cagggacagg cgtggacgac cctgcgggtg ccggagattc gggcgaatcg 36000  
atgatccagc aggcgcgct ggtggaggcg atcggtcagg ccgctcgggt cgggtccaca 36060  
tagccgcccg gcatccggcg gtagtagttc gcctcgaagc ctgacgggc atccggaagc 36120



cagcggggaa	gccgctggag	aggctcaccg	ccatccgctc	acctggcatc	tgcgggaccg	36180
ctgatcgcg	acggctcggg	gaagtgtcg	tcgaaccacg	agacgaccac	tgcgagctg	36240
gccagggcgg	cgggaaagt	agccaatccg	gagagcggat	gccaccgcac	tggcgtaccc	36300
gccgcgcgg	agctgtccc	gagtcgctcg	ccgaatgcga	acggaacgat	ctcgtcgtcc	36360
gtgctgtggt	agacgagcgt	ggggaccacc	gggccaccgt	tcctacctgc	gacgctttcg	36420
gccagtcgtg	cgcgccatcg	aggttgctcg	aaaaggccgg	aagtgtcgag	gaagtgcgtc	36480
agctcgcggc	cgaggaagcg	ggtgacgagc	tccggtgcac	cgagctcgcg	cacttgatca	36540
acggcgggtac	gacccgcttc	ggtgagaagc	tcgtcgaaatg	gcagatcggg	gtaggcagcg	36600
gcatgcccga	ccaggccggc	cagcaccggc	ccggtgaaca	ccccgtcatt	tcggtggatg	36660
atgtccagca	gatcgatcgg	caccgcacct	gcggccgcag	cgcggttcg	cagttcaggt	36720
gcgtaggtgg	ggtgcagttc	gccggcgaag	gccgacgctt	gccaccctg	cgcatagccc	36780
cagatgccga	ccgggcagtc	ggtcgtcagg	ccggagcccc	gtagccgttg	cgcagcgcgg	36840
gcggcatcga	gcatggcgtg	tccttgccgc	ctgccgacgg	tgtaggtgtg	ggttccagga	36900
gtaccgaggc	cttcgtagtc	ggtgatgacc	acggcccacc	cgcggtcgcg	ggccacggcg	36960
atcagctcgg	tctccggctc	ggttccgggt	cgaagcaggt	acgacggggc	aacttggtta	37020
ccgaggccgt	gggtgcccac	tgcgaaagt	atgatggggc	gatcttcgcg	cggccacggg	37080
atgttcggca	ccagaacggt	gccggagacg	gcgttcggca	tgccaagggc	ggagttggac	37140
cggtagagga	tttgccaggc	cttggtgcg	acgggttcgc	ccgtgccgcg	cagtgccgag	37200
acgggcccgg	ccctgaggag	cgtgcccggg	acacccggcg	gtagcggcgt	cggcggtcgg	37260
tagaagggat	catccgcggg	tgcccgcaga	tcgtcgccga	ccaggctggc	gtgctcggag	37320
gccatcagga	ctgcttcttt	cgagcctgca	ggagcatgaa	acctatgctt	tcctcgtttc	37380
tggcgtaatc	cggtatgttt	cggtattccg	caaccgcggc	gatcagctgt	gctgggtccc	37440
gtccgtgctt	cgcgcgatg	tctcccaagt	agcgttgctg	gtaggtgccg	acagccgcag	37500
gctcgacgcc	ggcgagctca	tcgagtttcc	ggagcaactc	gtcgacgtac	caggagacca	37560
tgcacctggt	ctgtgccgtg	aggtcgggtg	cttcgagaat	ctcgaacccg	gcttcgctga	37620
ccagcgccgt	gaagctgttc	aaggtatggg	cggtcgtgcc	cgtccaaacc	gccgcgtact	37680
cttcggggag	tcgaacccga	gtgatgatgt	ctccgaggac	gaaccggccg	ccgggttcca	37740
ggattcgggt	gacctcgcgg	atcgcggcgg	cctgggtccac	gatctgcacg	acggactgca	37800
tcgcccacgc	ggcctgaaag	aaaccgtccg	ggtagggcag	ctgggcgcgg	tcgactagat	37860
cgaactcaag	actgcgggcc	agtcgggtct	cgttggcgag	cctggtggcg	gcggcgagat	37920
gctgggcgtt	cacggtgatt	ccggtgactc	gaacgccgct	ggcgcatgcc	gcacggacta	37980
cggtgtgccc	attgccgcag	cccaggctga	acagggtgcg	tcggggacgg	agcgcgccct	38040
tgtcgatgaa	caggtcgggt	agttggtcgg	cagcatccga	ccacgggtgt	gcaccggcat	38100
ctcccgcata	ccgcgccgcc	cagtaaccgt	ggtgcagggg	acgcccggtc	gccaacgcac	38160
cgaagatgga	ctccacctga	tccgcggctg	gaaatgcctg	tgtgttcgcc	cctctgctgt	38220
tcactcgtcc	tccgcgctgt	tcacgtcggc	cagggtgcaat	atgtcgtcca	gactccttgg	38280
cacccaagca	ggaacgcgcg	cttcggcggt	gacgcctttc	tcagggaacg	cgatgttgtg	38340
gtaggtgtgg	aggccgacca	aattgcgttc	caggtagctc	ggctcgtacg	agcccgcatg	38400
cggctgctcc	tcgtgctgaa	cgccttccaa	caggttcctt	agcaggctga	ccgtgggtgc	38460
gggtgcggcc	gggcaactgc	cctgcccggc	gaatccggga	gcataggctg	tcacagatc	38520
ctcgatcacg	tatacgccac	cgtgcgcaa	ccgggggaac	agcgtttcca	gggatgtgcg	38580
caegtgtccg	ttgatgtggc	tgccatcgte	gatgatgatg	tcgaaaggte	cgtaactgtc	38640
gtcaacggcg	gccagctcct	cgggcttgct	ctggtcggcg	cggacgggtc	agagcctctg	38700
ctggctcagg	aaggacttgt	cgaaaacgtc	catcccgaac	acgaggccgc	ggtggaagta	38760
gcgcttccac	atcttcaggg	attcgccggc	gccaccgtcg	aagttgtagc	caccgacacc	38820
gatctccagg	atgcgcaccg	ggcgatcacg	gaactcggcg	aggtgtcgct	cgtatagcgg	38880
ggtgaaccag	tgcaggccgc	cccacttgct	cgtgcggtag	tgggaggcga	gcaagttgag	38940
gtcgggacgt	cggtgcccgc	agccggcgac	cactgcggag	atggcctgga	agccatcgga	39000

cagttccgac ggaccggg tcgaaccgga tgtggtggtt cggagg tgggtgctccg 39060  
 ggcgccgacg gccctgggag ctccctgggcc gaacaactcg gcgatgagat cggtgagctc 39120  
 gtaaccgatc cgcagcggga cgtctccgac cggtcgttgc tcggccttga tcagctcacc 39180  
 ggactgtagc gtcaggacga agtcaacggg ctccgctcgg tgggtgatct ggaccgagc 39240  
 ctccggtccgt tcgatgtcgg gggccgggtt cgcgcggaag aggatctcgt cgatcagcac 39300  
 ggggtgcgac ctggcgagtc cgagttcggg ggtcaggtcg gccaggctcg ccgactgga 39360  
 tccggcgggc aggatgatgc gttccacggg ttogatctcg tgcgttgtgg acatcgtgat 39420  
 gagctcctca tggctgaccg ggtgaaagcc gtgocggcgg tttgatcgac aggcctgtgt 39480  
 ggaagatgtt ctgcggatcc caccgcgctt tggcccgctg cagcccgggg tagttgtctt 39540  
 tgtagtacag gtcgtgccag gcaacaccgg aggtgttcca caatggatcg gccaaagtcgg 39600  
 tgtccgggta gttgatgtag gagccgtcga caccgggtacc tggcaccgga actccgccc 39660  
 tttcggcgta catctcgcgg tagaaaccgc gaatccaggt cagatgccgc tcgtcctcgg 39720  
 cgggctccga ccagttcgtg acgaacagcg ctttgagaac cgagtcgcgc tgagcgagtg 39780  
 cgggtggccga cggagccacg gcattcgcca taccgccgta accgagcagc aacagcgccg 39840  
 ccgagggtt gtcgtatccg tagacggtca gccgccggtg aaccgtgggt agttgagctt 39900  
 cggacagccc ggtgcgcaag taggcggctt tgaccttggg ccgttgcatg cccggttcgc 39960  
 cgccttcggc gatcgccccg gccacctggg tcgatcgcaa ccacggcagg gtttcccga -40020  
 gcccttcggc tggagtcacg ccgacctggg cgttgatcgc cgacagggtg tcggccaggg 40080  
 tgcgttcgcg gttcggatcc gtgccgtcca ggtgaacggt cagcgtgacg tagccagctt 40140  
 gccggtgtgc gcagacgagc gtgctgaaca acccgagttg cgtggattca ggcgcgctgt 40200  
 gctgctcgta ccaattgccg aagttctgta ggagcacggc gaatgactgc tctgtcagtt 40260  
 cgtgccacgg ccagtggaaac gatcggagca gcaactgtcg gggcgccgtt ggcaggagct 40320  
 ctgcggcgtc ggtgctgacc acgtccggcg ttcggagcca aaacctgggt acgatcccga 40380  
 agttgccgcc accgccaccg gtgtgcgccc accacaagtc gtgaccggcg cccgtggagt 40440  
 tccggtcggc ctcgacgatg tgcacttcac cggcctgggt gaccacgacg acctcgacgc 40500  
 cttgaaggta gtcgacgacc gaaccgaatc ggcgcgacag cgggccgtat ccccgccga 40560  
 ggatgtgccc gctgcgccc accccgggac atgcgcgggt cgggatcgtc acgccccagt 40620  
 tcttgaacag ggttcgggtac acctgcccga gggcggcgcc cgcctcgatc gcgaatgcc 40680  
 cgcgcgtgct gtcgtagtac acgcggttga gtcggagag gtcgacgagc actcggatcg 40740  
 ccgggtccgc aacgagattc tcgaagcagt gcccgccgct gcggaccctt accgcctgc 40800  
 cgggtgcgac ggcgtcggcg acggcgtgca cgacgtcttc ggcggagctg gcgatgtgga 40860  
 tgcgttcggg ttttcgggtg aaacgggggt tgtgcccagc gacgaggctc ggataacgag 40920  
 gatcgtcggg ctcgacgggtg atctctgttc ctgggggttc acgattcatg ggtgccgggt 40980  
 atggaattc gggcacggcc cctccttttc tgactggtec actttgttcg cccgcagccg 41040  
 agatcatcta cgcgtccggg tgattatctg tgtgtttcag ctcatagctg aaaccgggtc 41100  
 gcctccgccc gctctacttt gtggatcgat atcgcggtgc gcatggtgcc gtatgcgctg 41160  
 gaaccgaaaa ggtgatgact ttccatgagt gagatcgag ttgccccctg gtcggtgggtg 41220  
 gagcgtttgc tgcctcgccc ggggtcgggc ccggcgaagc tccaggaagc agtgcagggtg 41280  
 gccggactgg acgcggtggc cgacgccatc gtcgacgaac tcgtcgtacg ctgcgatccg 41340  
 ctgtcgttgg acgagtcggt gcgaatcggc ctggagatca cttctggcgc tcagctggtc 41400  
 cggagaaccg ttgagctcga tcacgcaggc ctgcggctcg cggcggtcgc cgaagcagct 41460  
 gctgttctcc ggttcgaagc ggtggatctg ctggaagggc tcttcggccc ggttgacggc 41520  
 aggcggcaca acagccgtga agtccgctgg tcggacagca tgacgcagtt ctgcccgcac 41580  
 cagggcctcg ccggcgcgca gcgcctgctg gcgttcggga acagggtgtc caccgcgggtg 41640  
 caccgcgtgc tggccgcagc cgccaccagg cgcgcggacc tcggtgcgct ggcagtcgcc 41700  
 tacggatccg acaaattggc ggacctgcac tggtagacag aacactacga gcaccacttc 41760  
 tcccgattcc aggatgcccc ggtgcgagtg ttggaaatag gaatcggtgg ttatcacgca 41820  
 cccgaactcg gtggtgcttc gctgcgcatg tggcagcggg acttcggcg aggtctcgtt 41880

tacgggctgg	acatttttcca	gaaagccggg	aacgaagggc	accgagtgcg	aaagctgcga	41940
ggtgaccaga	gcgatgcgga	attcctggaa	gacatggcgg	ggaagatcgg	gccgttcgac	42000
attgtcatcg	acgacggcag	ccatgtcaac	gaccacgtca	agaaatcctt	ccaatccctg	42060
tttccgcacg	tccgcccagg	tggtttgtac	gtcatcgagg	atctccagac	ggcgtactgg	42120
cccggctacg	gcggtcgcca	tggggaaccc	gcggcccagc	gcacctcgat	cgacatgctc	42180
aaagaactga	tcgacggcct	gcattatcag	gagcgcgcaat	cgcggtgcgg	gaccgagccc	42240
tcctacacgg	aacggaacgt	ggcgccctg	cacttctacc	acaacctggt	attcgtggag	42300
aaagggctca	acgctgagcc	tgcgcgcgg	gggttcgtgc	cccggcaagc	gctcggcgctc	42360
gaggacggct	gagccgttca	ccagctgcgg	cgccagtagg	cgcccgtgcc	gtcgatgtcg	42420
tggatgggtt	ccgtgatccc	gagttccgg	cggaacccct	tcaccgcgtc	ctggcaggac	42480
ggcagaaaat	agtcgtcgat	gatgacgaat	ccgcccggcg	agagcttcgg	gtacaggttc	42540
cgcaatgagt	ccattgtgga	ttcgtagagg	tcgcccgtcg	gtcgtagcac	ggcgagttcc	42600
tggatggggg	cggtgggcaa	ggtgtcccgg	aaccagccgg	ggaggaacct	gacctgttcg	42660
tcgagcagcc	cgtagcgggc	gaagtctctg	cggacgggtct	caagcgatac	gccaaagcacg	42720
tcgttgact	cgtgcagcgc	catagcctgg	tccgcttggg	ggtcttgccg	agagctttcc	42780
ggcattccct	ggaaggaatc	cactaccag	acggtacgtc	cggtatctcc	gaatgcctgg	42840
agaaccgcgc	gcatgaagat	gcatgcgccg	ccccgccaga	caccgggtctc	ggcgaaatcc	42900
ccgggaacac	cgtctgcgag	cacggcttcc	acgcagtgtc	ggaggttgtc	cagccgctcc	42960
agaccgatca	tcgtgtgcgc	gacagttggc	cagtcctgtc	ctttggcccg	agcggcctgc	43020
ctgtagtccg	tgttgctctg	ccaggcgttc	ggatgcggcc	gatcaactgt	aatcgtgttg	43080
gtgagtacct	tcttgagcag	gtccaggtac	agcgcgttct	gggagggcat	cggttctccg	43140
gatccagctg	ttctcgggtg	actagttcat	caggcacgga	tggccgcagt	gttctccagc	43200
gtccgcacca	gcgcggcggg	atggggcatg	gccgtgatct	cgtcgctgag	tttgattgcc	43260
gcagcagcca	agccggtgtc	gccgagcacc	gttgcgattg	agtcggtgaa	ctgttcgtgg	43320
tcggactggg	cctgctcacc	cggcaagcag	atgccgcccc	cggcagcggc	gaggttgccg	43380
gcgtagtcga	actggtcgaa	gtactgggga	agcacgagtt	gcgggatgcc	gagtcgggtc	43440
gcggtgaatg	ccgttcccga	gccgcccgcg	cagatgacca	gctcgcaggt	acgcaggaac	43500
aggttgagcg	ggaccgattc	ggcgatccgg	gcgttgctcg	gtaggtcggt	gagaagtgcc	43560
cggtgctcgg	ggggaacggc	gatcacggcc	tcgacgccgg	gcaactcggg	ggcagccgct	43620
actgcgcgca	gcagcggagc	cggcccgggtg	gcgttcagca	ccatgcggcc	catgcagatg	43680
cagacccgcc	gtgctgaggt	gcgcgcgcgg	ccccatgccg	ggaatgcgcc	gcttccgttg	43740
tacggcacgt	actggaccgg	tgcgccttgc	ggcgcgtcgc	ttgcttgca	gctcggcgga	43800
cagggatcga	ggatgagctc	gggagtgggc	aggccgggtca	gtccgtgggtg	ccggcacacc	43860
gggtcaagca	actcgtgggc	tcgatcgctg	aaggggcctg	cgggtggggtc	gactccccag	43920
cgggtgcagca	cgaccggcag	gtcgagcaat	ccgccgagca	cccggccgat	cagcgcgcag	43980
acgtcgacca	acagcactga	cggtcgccag	gcctcggcca	gtcgaaggta	ttcggggagc	44040
tgatcgagcg	agctttgcgc	gacattggac	gcggtctgct	cccacagttg	ccggcctgcc	44100
tcggtgtcgc	gctgaccgaa	cgcgcgattg	ggaaagcgca	gctgcgtggg	tccacccgta	44160
tcgccggtcc	tgtcgttccc	gcggatcccg	gccgtgggtg	gacctgcacc	atgcgcggtc	44220
gcctgcagct	ctggtgggtg	ggcgatcagg	acctcgtgcc	cggatgcttg	cagcgcgccag	44280
cacagcggca	ccattgccat	gagatgcgtc	ggatagggca	agggaaacgac	gagtaacgcgc	44340
atacttcgga	ccccagttct	tttcccccca	ttagcgcagc	agccctact	cccatgtggc	44400
aggatttgga	aaatgcgctg	cgtatgtcga	tcgccgttga	cgtccaaacgg	acttccggcg	44460
gcaacaatag	tgtgtcacgg	caggaatgtc	acgcgaccat	cgaagatctt	tgggtgcgcc	44520
cacctggttt	cacgcgaacg	agtgaatgc	gcgagctccg	ctcgatcggg	gtgggcccga	44580
cctgtacggg	gatcacctgt	ggttctgcgg	ggattcatgg	ggaagatttg	cgctggctgt	44640
ttgcctcctg	gccgataggt	tatagtcggg	accgcgcgat	gcggcggtaa	ccgcgaatta	44700
actgacggct	agtttgccgt	cttttctctc	tgtgtgtttc	ctgctcgggt	ccagaaaatt	44760

acgagaaggt gaacgttga gagatcaggc ataccggtgt tgccaggagg cgcaccaaca 44820  
 tcgcagcagg ttgggcagat gtatgacctg gtcacgccgt tgetgaactc ggtcgcgggc 44880  
 ggcccctgcg ccatccacca cggctactgg gagaacgacg ggccgggcttc ctggcagcag 44940  
 gccgccgacc ggctcaccga cttgtgcgc gaacggaccg tgctcgatgg cggcggttcga 45000  
 ctgctcgatg tgggggtgcg taccggacaa ccagcgctgc gcgtcgcgcg cgacaacgcg 45060  
 atccagatca ccggcatcac cgtcagccag gtgcaagtgg ccatcgccgc tgattgcgca 45120  
 cgcgaacgcg gactaagcca ccgggtggac ttctcgtgcg tcgatgccat gtcctgccc 45180  
 taccgggaca atgctttcga cgccgcctgg gccatgcagt cgctgttgga gatgtccgaa 45240  
 ccggaccgtg ccatccggga aatccttcga gtactcaaac ccgggtggcat cctcggcgtc 45300  
 accgaggtcg tcaaacgaga agcggggcgc gggatgccgg tgtccgggga caggtggccg 45360  
 accggccttc ggatctgcct ggctgagcaa cttctggaat cgctgcgtgc agcggggctc 45420  
 gagatcctcg attgggagga cgtgtcgtcg aggaccgggt acttcatgcc gcagttcgc 45480  
 gaagagctcg ctgcgcacca gcacgggatc gcggacaggt acgggccggc tgtcgcggc 45540  
 tgggcccgcg cggctctcga ttatgagaaa tatgccacg acatgggcta tgcgattctg 45600  
 acggcgccga agccggctcg ctgagggcgc gccgcaattc gatgacgttc atgcgccgtg 45660  
 tcggggaatc gccggtggcg gcgccagcag aggtgaact tactgggtgt gtgtccagga 45720  
 atcggagggg cagtaccgaa tgagcgaagc cgggaacctg atagccgtca tcggactgtc 45780  
 ctgccgccta cccagggcgc ctgaccggc ttcttctggt cggttgctgc gcaccggaac 45840  
 ggacgccatc accacggtcc cggaagggcg gtggggcgac ccgttgccctg gtcgggatgc 45900  
 gcccaagggc ccggaatggg gtggtttcct ggctgatgtc gactgcttcg atcccagatt 45960  
 cttcgggatc tcgccgcgag aagcggcagc cgtggatccc cagcagaggc tggctctgga 46020  
 gctcgcctgg gaggcactcg aagacgccg tatccccgcc ggccagctgc gcggtactgc 46080  
 cgccgggtgtg ttcattgggg cgatctctga cgactacgcc gccctgctgc gcgagagccc 46140  
 gccggaagtg gctgcgcagt accgcctcac cggcaccat cgaagcctga tcgccaaccg 46200  
 cgtgtcctat gtgctcgcc tcgcgcgggc aagcctgacg gtggattcag gtcagtcctc 46260  
 gtccctggtc ggctgcac tcgccagcga gacgtgcga cggggtgagt gcacgatgc 46320  
 actcgcggc ggctgaacc tcaacctggc cgccgagagc aacagcgctc tgatggactt 46380  
 cggcgcgctc tccccggacg gtgctgctt cacttcgat gtgcgggga acggttacgt 46440  
 ccgtggtgag ggccggggcc ttgtcgtgct gaagaaggcc gatcaggcgc accccgatgg 46500  
 cgaccggatc tactgcctca tccgcggcag ccgggtcaac aacgatggg gcggtgccg 46560  
 gctcaccgtt ccggcgggcg acgcccaggc ggagctgctg cgccaggcat accggaacgc 46620  
 gggcgctgac ccggccgcg tgcatatgt cgagctccac ggacgcgga ccagggtcgg 46680  
 ggatcccgtc gaagcagcag ccctcggagc tgcctgggg gcggcgagac ggcccgga 46740  
 cgagctgct gtggggtcgg cgaagaccaa cgtcgccat ctggaagcag cggcgggcgt 46800  
 caccgggttg ctgaagacc cactcagcat ctggcaccgc gaactgccgc cgagtcttca 46860  
 tttcaccgcc cccaaccgg aaatcccgt ggacgaattg aacctacgc tccagcgtga 46920  
 tctgcggccg tggccggaga gcgagggggc gctgctggcc ggctcagcg ccttcggaat 46980  
 gggaggcacg aactgccacc tgggtgcttc cgccagctcc cgggtggagc gacggcgag 47040  
 tggaccgct gaggcgacca tgccgtgggt cttgtcggcc agaacaccgg tcgattgag 47100  
 tgccgaggcg gcgcgcttg acacgcacct caatacgcc ggtcaaagtc cgttgagcgt 47160  
 cgctactca ctggcgacca ctgatccgc gctgcgcac cgggcccgcg tggtcgga 47220  


---

 cgaagaaacg aaactgctg eeggggtgaa ggeectegat gaaggagagc aeggeccac 47280  
 gctgtgccac ggccgactt ccggcgagcg ggcagcggtc ttctctttc ccggacagg 47340  
 cagccagtgg atcgggatgg gtaggcagct gtcgaaacc tccaggttt tcgcgcgctc 47400  
 gatgtcggac tgcgccgac cattggcgcc gcacctggat tggctcctgc tggatgtgct 47460  
 gcgcaacgcg gccggcgct cgacattga ccacgacgat gtcgtccagc ccgcgctgtt 47520  
 cgccatcatg gtctcgtcg cggagctctg gcgttcgtg ggctgcgtc cgggtggcgt 47580  
 cgtcgggcac tcgaggggg agatcgccg gcctgcgtc gccggggccc tgtccgtccg 47640

cgatgccgcc aggggtggtg cggtgcgcag caggcttctg acggcgctgg ccggcagtg 47700  
 cgcgatggcc tcgttgacgc atcccgcga agaggtgcgg caaatcctgt tgccctggcg 47760  
 cgatcggatc ggcgtggcgg ggggtgaacgg accgtcgtcg accctggtgt caggggaccg 47820  
 ggaggcgatg gcggaactgc tggccgagtg cgcagaccga gagctccgga tgcgcggat 47880  
 tcccgttgaa tacgcctccc attcgctca catcgagggt gtccgggatg agctgctggg 47940  
 gctgttggcg ccggtcgaac ccaggacggg aagcatccc atctattcga cgacgaccgg 48000  
 ggacctgctg gaccggccga tggacgccga ctactggtac cgcaaccttc gtcaaccggt 48060  
 gctgttcgaa gcggccgtcg aggccctgtt gaagcggggg tacgacgcat tcatcgagat 48120  
 cagcccacac ccggtgctga ctgcgaacat ccaggaaacc gccgtgcgag cagggcgagg 48180  
 ggtagtggcg ctccggacac tccgcccgcg cgaaggtggc atgcggcagg cgctgacgtc 48240  
 gctggccaga gcacacgtac acggagtggc cgcggactgg cagcgggtct tcgccggtac 48300  
 cggggcgag cgggtcgacc tgccgacgta cgcctttcag cgacagcgct actggctgga 48360  
 cgcgaagctt cccgacgtcg ccatgcccga gagcgacgtg tcgacggcgt tgcgggaaaa 48420  
 gctgcggctc tcgccgaggg cggacgtgga ctgcagacc ctacgatga tccgggcaca 48480  
 ggcagccgtg gtccctcgcc actccgatcc gaaagaggtg gaccggatc ggacgttcaa 48540  
 ggacctgggc ttcgattcct cgatggtggt cgagctgtgc gaccgcctaa acgccgccac 48600  
 aggtctgcga ctgcaccca gcgtcgtttt cgactgtcct acgccggaca agctcggccg 48660  
 ccaggtacgg acgttggtgt tgggcgagcc ggctcccatg acgtcacacc ggccggactc 48720  
 cgatcgggac gaggctatcg ccgtgatcgg gatgggctgt cggtttccgg gtggggtgtc 48780  
 ctgcgccgag gattgtggtc agttggtcgc cgctgggcgg gacgtcgtgt ccgagttccc 48840  
 ggctgaccga ggttgggacc tggagcgtgc ggggacatcg cacgtgcgcg ccggcggggt 48900  
 cttgcatggc gcccgggatt ttgacccgg gttcttcgg atttcgccgc gcgaggcgtt 48960  
 ggcgatggat ccacagcagc ggttgctgct ggaaatcgcc tgggaagcag tcgaacgagg 49020  
 cgggatcaac ccgagcacc tgcacggaag tcaaaccggg gtcttcgtcg gcgcgacctc 49080  
 cctggactac gggccacgcc tgcacgaagc gtccgaggag gcggccgggt acgtgctcac 49140  
 cggcagcacc acgagtgtgg cgtcgggtcg ggttgcgat tcgttcgggt tcgagggcc 49200  
 tgcggtgacg gtggatacgg cgtgttcgtc gtcgttggtg gccctgcatt tggcgtgtca 49260  
 gtcgttgctg tcgggtgagt gtgatctggc gttggccggg ggtgtgaccg tgatggccac 49320  
 gccggggatg ttcgtggagt tttcgcggca gcgtggtttg gcgccggatg ggcggtgcaa 49380  
 gtcgttcgcg gaggccgcgc acggcaccgg ctggtccgag ggtgctggcc tggttctact 49440  
 ggagcgggtt tcggatgccc ggccggaatg gcatgaggtg ctggcgggtt ttcgtggtag 49500  
 tgcggtgaat caggacggtg cgtcgaatgg tttgaccgc ccgaatggtt cgtcgcagca 49560  
 gcgggtgatt gcccaggcat tggcgagtgc ggggttgctg gtgtccgatg tggatgctgt 49620  
 ggaggcgcat gggacgggca cgcggcttgg tgatccgatc gaggcgcagg cgctgatcgc 49680  
 cacctacggc cagggccggc ttccggaacg gccattgtgg ttgggctcga tgaagtcaaa 49740  
 catcggtcac gcgcaggcag ctgcggggat agccggcgct atgaagatgg tgatggcgat 49800  
 gcggcacggg cagctaccgc gcacgttgca cgtggatgag ccgacttctg gggtggtatt 49860  
 gtcggcgggg acggttcaac tccttacgga gaacacgccc tggcccgga gtggtcgtgt 49920  
 tcgtcgggtg ggggtgtcgt cgttcgggat cagtgggtact aacgcgcacg tcatcctcga 49980  
 acagcccccg ggagtgcga 50000

<210> 2

<211> 50000

<212> DNA

<213> *Saccharopolyspora spinosa*

<400> 2

gtcagtctgc ggggccggtc tcgggctctg tcgtggatgt tccggtggtg ccgtggatgg 60  
 tgtcgggcaa aacacccgaa gcgctatccg cgcaggcaac ggcgttgatg acctatctgg 120  
 acgagcgacc tgatgtctcc tcgctggatg ttgggtactc gctggcggtg acacggtcgg 180  
 cgctggatga gcgagcgggtg gtgctgggggt cggaccgtga aacgttggtg tgcgggtgta 240  
 aagcgctgtc tgccgggtcat gaggttctg ggttggtgac cggatctgtg ggggctgggg 300  
 gccgcacggg gtttgtgttt tccgggtcagg gtggtcagtg gctggggatg ggccgggggc 360  
 tttaccgggc ttttccgggtg ttgctgctg cctttgacga agcttgtgcc gagctggatg 420  
 cgcacatctggg ccaggaaatc ggggttcggg aggtggtgtc cggttcggat gcgcagttgc 480  
 tggatcggac gttgtggggc cagtcggggt tgttcgcgtt gcaggtgggc ttgctgaagt 540  
 tgctggattc gtgggggggt cggccgagtg tgggttggtg gcattcgggtg ggcgagttgg 600  
 cggcgccggt cgcgccgggt gtggtgtcgt tgtcgggtgc ggctcgggtg gtggcgggtc 660  
 gtgcccgggt gatgcaggcg ttgccgtctg gcgggtgggt gctggcggtg cctgctggtg 720  
 aggagctgtt gtggtcgttg ttggccgac aggggtgatc tgtggggatc gccgcgggtc 780  
 acgctgcggg ttcgggtggt ctctctggtg atcgggatgt gctcgatgac cttgccgggtc 840  
 ggctggacgg gcaagggatc cggtcgaggt ggttgcgggt gtcgcatgcg tttcattcgt 900  
 atcggatgga tccgatgctg gcggagttcg ccgaattggc acgaaccgtg gattaccggc 960  
 gttgtgaagt gccgatcgtg tcgacctga ccggagacct cgatgacgtt ggcaggatga -1020  
 gcggggccga ctactgggtg cgtcagggtc gagagccggt ccgcttcgcc gacggtgtcc 1080  
 aggcgctggt cgagcacgat gtggccaccg ttgtcgagct cggtcgggac gggcggtgtg 1140  
 cggcgctgat ccaggaaatg gtgcgccat ccgatcacgc cggcgcggtg agcgcggtcc 1200  
 cggcgatgcg caggaaaccg gacgagggcg agaaggtgat gacggccctg gcacacgtcc 1260  
 acgtacgtgg tgggtgcgtg gactggcggt cgttcttcgc cgtacaagg gcgaagcaaa 1320  
 tcgagctgcc cacctacgcc ttccaacgac agcgggtactg gctgaacgcg ctgctgtaat 1380  
 cttccgccgg cgacatgggc aggcgtgtcg aagcgaagtt ctggggcgcc gtcgagcacg 1440  
 aagatgtgga atcgcttgca cgcgtattgg gcattgtgga cgacggcgct gctgtggatt 1500  
 ccctgagaag cgccttcctg gtgttgcccg gttggcagcg aaccgcacc accgagtcca 1560  
 ttatggatca gcggtgttac cgaattgggt ggcggcaggt agccggactc ccgccgatgg 1620  
 gaactgtttt cggtaacctg ctggtcttcg cgcctcatgg ctggtccagc gaaccggagg 1680  
 tgggtggactg cgttacggca ctgcgggcac gtggtgctc ggtggtgttg gtggaagctg 1740  
 atcccgaacc gacctcttc ggcgaccggg tacgaacct gtgttcgggc cttccggatc 1800  
 ttgttggcgt gttgtcaatg ttgtgcttg aagaatcgggt ccttcggga tttctgcgg 1860  
 tgtcacgggg ttttgcgttg accgtggagt tgggtcgggt tttgcgggca gctggtgca 1920  
 ctgcccgggt gtggttgctg acgtgtggtg gcgtgtcgggt gggagatgta ccggttcgtc 1980  
 cagcgcaggc cctggcggtg ggggtggggc gtgttggtgg gttggagcat ccggactgggt 2040  
 ggggcggcct gatcgatatt ccggtcttgt tcgacgaaga cgctcaagag cggttgtcga 2100  
 ttgtgctggc aggtctcgat gaggacgagg tcgcatccg tctgacggc atgttcgcgc 2160  
 gtcggttggg acgccacact gtctcagctg atgtgaagaa ggcgtggcgc ccaggggat 2220  
 cgggtgctggg gacgggcggc acgggtggtt tgggggcgca cgttgctcgc tggctggccg 2280  
 acgccggagc cgaacatgtg gcgatggtga gtcgacggcg cgagcaggca ccgagtgtg 2340  
 agaagttgcg gacggaactg gaggatctgg gtaccgggt gtcgacgtg tcatgcatg 2400  
 tgaccgatcg cgaggcgtc gccgaagtgc tgaaagccct tccggctgaa aaccggttga 2460  
~~cgcggtagt geatgggga ggcgtgatg agactggtg tgcggggga atgagcctgg 2520~~  
 ctgatttcga tcacgtgttg tccgcaaagg tggccgggtc cgcgaatctg gatgccttgt 2580  
 tggccgatgt ggaattggac gcgttcgtct tgttctcctc ggtgtcagga gtttggggcg 2640  
 ctgggggaca cggggcttac gcagcggcga atgcctatct ggatgcgctc gcggaacagc 2700  
 gtcggtcgcg agggctggtc gcgactgcgg tggcctgggg gccgtgggccc ggcgagggca 2760  
 tggcctccgg agaaacagga gaccagctgc gccgatacgg cctttcccca atggctccgc 2820  
 agcacgccat cgcgggaatc cggcaggccg tggaacagga cgaaatttcc ctggtagtgg 2880

ccgatgtcga ttgggcacgt ttcagcgcg gatttgcctggc ggctaggccg cgcccgctgc 2940  
tgaacgaact ggccgaggtc aaggaaactcc tcgtcgatgc ccagcccag gcgggagtc 3000  
ttgccgacgc gtogttggaa tggcggcagc gattgtccgc ggcaccgagg ccgacacagg 3060  
aacagctgat cctggagctg gtacgcggcg aaaccgctct ggtgctggga cccccgggg 3120  
cagcggccgt tgcacgga cgagccttca aggacagcgg attcgactcg caggccgagg 3180  
tcgaactccg cgttcggctc aatcgagcta ccggcctcca gttgccatcg acaattatct 3240  
tcagccatcc cagccctgcg gaactggctg cggagctgcg ggcgaggctt cttcccaggt 3300  
ccgcaggagc aggcattccc gaggaggagc aggcgcgaat cagagcggca ctgacgtcga 3360  
tcccgttccc ggccttgccg gaggcaggct tggtagtcc gctgctcgca cttgccggac 3420  
accgggtcga ctccggatc tctcggagc atgcggccgc gacctcgatc gatgcgatgg 3480  
atgtagccgg cctcgtcgaa gcagcgtgg gcgaacgcga gtcctgagac cgccgacctg 3540  
ggagatgacg gtgaccacca gttacgaaga agttgtcgag gcaactgcgag catcgctcaa 3600  
ggagaacgaa cgcctccggc gcggcaggga tcggttctcc gcggagaagg acgatcccat 3660  
cgcgatcgtg gcgatgagtt gtcgttatcc cggtcaggte tctcggccg aggcctgtg 3720  
gcaactggct gccggcggtg tggacgcgat ctccgaagtt ccgggggatc gcggatggga 3780  
cctggatggc gtgttcgttc cggactccga tcgtcctggc acgtcgtatg cctgcgcggg 3840  
cggttttctt cagggcgtgt cggagttcga cgcgggttct ttcgggattt cgccgcgtga 3900  
ggcgttggcg atggatccgc agcagcgggt gctgctggaa gtcgcgtggg aggtcttcga 3960  
gcggttggcg ctggagcagc ggtcgacacg cggttcccg gttggcgtgt tcgtcggcac 4020  
caatggccag gactacgcgt cgtggttgcg gacgcgcccg cctgcgggtg caggatcatgt 4080  
gctgacgggc ggtgcggcag cggttcttct gggccgggtt gcgtattcgt tcgggttcga 4140  
gggtcctcgc gtgacggtg atacggcgtg ttcgtcgtcg ttggtggcgt tgcacctggc 4200  
ggggcaagca ctgcggggcg gtgagtgcga ccttgccctt gccgggtggc tcacggtgat 4260  
gtcgacgccg aagggtgttc tggagttct cccccaacgg ggtctcgcgc cggatgggcg 4320  
gtgcaagtcg ttcgcggcg gtgcggatgg cactggatgg ggtgagggtg ccggactgtt 4380  
gttgcgtggg cggttgtcgg atgcccggcg gaatgggcat gaggtgctgg cggttgttcg 4440  
tggtagtgcg gtgaatcagg acggtgcgtc gaatggtttg accgcgccga atggttcgtc 4500  
gcagcagcgg gtgattaccc aggcgttggc gagtgcgggg ttgtcgggtg ccgatgtgga 4560  
tgctgtggag gcgcattgga cgggcacgcg gcttggtgat ccgatcgagg cgcaggcgct 4620  
gatcgccacc tacggccgtg atcgtgatcc tggccggccg ttgtggttg ggtcgggtcaa 4680  
gtcgaacatc ggtcatacgc aagcggcggc ggggtgtggc ggtgtgatca agatggtgat 4740  
ggcgatgcgg caggggcagc tgccacgcac gttgcacgtg gaatcgccgt cgccggagggt 4800  
ggattggctg gcggggacgg ttcaactcct tacggagaac acgccctggc ccaggagtgg 4860  
tcgtgttcgt cgggtggggg tgcgtcgtt cgggatcagt ggtactaacg cgcacgtcat 4920  
cctcgaacag cccccgggag tgccgagtcg gtctgcgggg ccgggttcgg gttctgtcgt 4980  
ggatgttccg gtggtgccgt ggatggtgtc gggcaaaaca cccgaagcgc tatccgcgca 5040  
ggcaacggcg ttgatgacct atctggacga gcgacctgat gtctcctcgc tggatgttg 5100  
gtactcgtg gcgttgacac ggtcggcgt ggatgagcga gcgggtgtgc tggggtcgga 5160  
ccgtgaaacg ttgttgtgc gtgtgaaagc gctgtctgcc ggtcatgagg cttctgggtt 5220  
ggtgaccgga tctgtgggg ctgggggccc catcgggttt gtgttttccg gtcagggtgg 5280  
tcagtggctg gggatgggcc gggggcttta ccgggctttt ccgggtgttc ctgctgcctt 5340  
tgaegaaget tgtgeegage tggatgaaea tetggggccg gaaatcgggg ttcgggagggt 5400  
ggtgtccggg tcggatgcgc agttgctgga tcggacgttg tgggcgcagt cgggtttgtt 5460  
cgcgttgacg gtgggcttgc tgaagttgct ggattcgtgg ggggttcggc cgagtgtgg 5520  
gttggggcat tcgggtggcg agttggcggc ggcgttcgcg gcgggtgtgg tgcgttgtc 5580  
gggtgcggct cgggtggtg cgggtcgtgc ccggttgatg caggcgttgc cgtctggcg 5640  
tgggatgctg gcgggtgcct ctggtgagga gctgttgtg tcgttgttg ccgatcagg 5700  
tgatcgtgtg gggatcgcc cggatcaacgc tgcggggtcg gtggtgctct ctggtgatc 5760

ggatgtgctc gatgaccg cgggtcggct ggacgggcaa gggatcgggt cgaggtggtt 5820  
 ggggtgctc catgcgtttc attcgtatcg gatggatccg atgctggcgg agttcgccga 5880  
 attggcacga accgtggatt accggcggtg tgaagtgcg atcgtgtcga ccttgaccgg 5940  
 agacctcgat gacgctggca ggatgagcgg gcccgactac tgggtgctc aggtgcgaga 6000  
 gccggtccgc ttgcgcgacg gtgtccaggc gctggtcgag cacgatgtgg ccactgttgt 6060  
 cgagctcggg ccggacgggg cgttgtcggc gctgatccag gaatgtgtcg ccgcatccga 6120  
 tcacgccggg cggctgagcg cgggtcccgc gatgcgcagg aaccaggacg aggcgcagaa 6180  
 ggtgatgacg gccctggcac acgtccacgt acgtgggtgt gcggtggact ggcggtcgtt 6240  
 cttgcgggt acgggagcga aacaaatcga gctgccacc tacgccttcc aacgacagcg 6300  
 gtactggctg gtgccatcgg attccgggtga tgtgacagg ggcggtctgg ccggggcgga 6360  
 gcatccgctg ttgggtgctg tgggtccggg cgcgggtgtg gacgaggtgt tgctgaccgg 6420  
 caggatttcg gtgcggacgc atccgtggct ggccgaacac cgggtgctgg gtgaagtgat 6480  
 cgttgccggc accgcgttgc tggagatcgc cttgcacgcg ggggaacgtc ttggttgtga 6540  
 acgggtggaa gagctcacc tggagcacc gctggctctg ccggagcgcg gggcgatcca 6600  
 ggttcagctg cgagtggcg cggccgagaa ttccggacgc aggcgatgg cgtgtattc 6660  
 acgccccgaa ggggcggcgg agcatgactg gacgcggcac gccacgggcc ggttggcgcc 6720  
 aggcgcggc gaggcggctg gagacctggc cgactggccg gctcctggcg cgtgcgggt 6780  
 cgacctcgac gaattctatc gggacctcgc agagcttggg ctggagtacg gccgatctt 6840  
 ccaagggctc aaggcggcct ggcggcaagg ggacgaggtg tacgccgaag ccgcgctgcc 6900  
 gggaaacggaa gattctggtt tcggggtgca tccggcactg ctggacgcgg ctctgcacgc 6960  
 aacggctgtc cgagacatgg atgacgcacg cttgccgttc cagtgggaag gtgtgtccct 7020  
 gcacgccaa ggcgcgcgg ctttgcgggt ccgcgtggtc ccggtggtg acgatgcaa 7080  
 gtccctgctg gtttgtgatg gcaccggtcg accggtgatc tcggtggacc gactcgtatt 7140  
 gcggtcggct gcggcccggc ggaccggtgc gcgccgacag gcccatcaag ctcggttgta 7200  
 ccggttgagc tggccaacgg ttcaactgcc gacatccgct cagccaccgt cctgcgtgct 7260  
 tctcggcacc tcagaagtgt ccgctgacat acaggtgtat ccggacctcc ggtcgttgac 7320  
 ggctgcgtt gatgccggtg ccgaaccacc cggcgtcgtc atcgaccca cgcggccgg 7380  
 cgggtggacga acagcggatg tccgggagac gactcggcat gactcgacc tggtaacaag 7440  
 ctggctttcc gatcagcgac tcaacgaatc ccgattgtc ctggtgacac agggagcagt 7500  
 ggccgtggag ccgggcgaac ccgtgaccga tctggcgcag gccgcgctc ggggactgct 7560  
 gcggtcgacg cagaccgaac accctgatcg cttcgtctc gtcgatgtgc ctgagccgc 7620  
 gcaactctc cccgcgctgc cgggggtgct ggcctgcggc gaacctcagc tcgcttgcg 7680  
 acgtggcggc gtcctatgcg ccagactggc tggactggc agcgatgacg tcctgccgt 7740  
 gccggacggc accgggtggc gattggaggc cacgcgcccg ggaagcctgg atgggttggc 7800  
 attggtggac gaaccgacgg ccacggcacc gctgggtgac ggtgaggtca ggattgcgat 7860  
 gcgcgcggcc ggggtgaact tccgggatgc gtcctatcg ctcggtatgt atcccggtgt 7920  
 ggcacgctg ggcagtgagg gcgcgggggt cgtggtggag accggcccg gcgtcaccgg 7980  
 cctggcacc gccgaccgcg tgatgggaat gatccgaag gcgttcgggc cgtcgcgggt 8040  
 cgccgacct cgcatggtga cgaggattcc cgtggttg agcttcgcgc gggccgcac 8100  
 ggtgccgat gtctttctca ccgctacta cgcgtggtt gatctgcgc ggttgagacc 8160  
 aggggagtcg ttgctggttc attcgccgc cgggtgggtg gggatggccg cgatccaaact 8220  
~~cgccaggaac ctegggtgag aggtgtacgc caccgctagc gaggacaagt ggcaagccgt 8280~~  
 ggagctgagc cgagaacacc tcgcttcgtc gcggacgtgc gatttcgagc agcagttcct 8340  
 cggggcaacc ggcggacgcg gcgtcgacgt cgtgctcaac tccctcgccg gggagttcgc 8400  
 cgatgcgtct ctgcgaatgc tgccgcgcgg tggcgtttc ctggagttgg ggaagacgga 8460  
 tggtcgtgac cccgtcgagg tcgccgatgc gcatccgggc gtgtcttacc aggtttcga 8520  
 taccgtagag gcaggcccg agcgaatcgg cgagatgctt cacgagctgg tggagttgtt 8580  
 cgagggacgc gtgctggagc ccctgcctgt cacggcttgg gacgttcggc aggcgccga 8640



ggcgctacgg cacctgagcc aagcgcgcca tgtgggaaag ctggtgctca ccatgcctcc 8700  
 ggtgtgggac gccgcaggca cggttctggt taccggcgga acgggagcac ttggcgagca 8760  
 ggtcgcccg cactcgtga tcgagcgcg ggtgcgaaac ctggtcctcg tcagcaggcg 8820  
 cggccccga gccagtggcg ctgctgagct cgtggcgcaa ctgacggcct acgggtgccga 8880  
 ggtttccttg caggcttgcg atgtcgccga tcgtgagacc ttggcgaagg tgcttgccag 8940  
 catcccgga gagcatccgt tgaccgccgt ggtgcacgcy gctggtgttc tcgacgacgg 9000  
 agtgtccgaa tcgctcaccg tggagcggtt ggaccagggt ctgcgcccga aggtcgtatg 9060  
 cgcgcggaat ctgctcgagc tgatcgaccc ggacgtggcc ctctgtgtgt tctcgtcggt 9120  
 gtcgggtgtg ctcggcagcg gtgggcaggg taactacgcy gcggccaact ccttcctcga 9180  
 cgcattggcg cagcaaaggc agtcgcgcy cctaccgagc agatcattgg cctggggggc 9240  
 ctgggcggaa catggcatgg ccagcacctt gcgcgaagcc gagcaggatc gattggcgcy 9300  
 atctgggttg ctgccgatct cgaccgagga ggggttgctc cagttcgacg ccgcgtgcy 9360  
 cggcgcgcat accgtggttg cgcgggttcg attcagccgc ttgtccgacg ggaacgcgat 9420  
 caagtctctc gtcttgcaag gtttggtcgg gccgcacgc gtcaacaaag cggcgactgc 9480  
 ggatgatgcc gagagcctcc ggaacgggtt gggacgcttg ccggatgcag aacaacatcg 9540  
 gattctgctg gacctcgctc gcatgcatgt ggcggcagtg ctcggtatcg ccggttctca 9600  
 ggagatcacc gcggacggca cgttcaaggc gctgggcttc gactcggtga ccgtggtcga 9660  
 gttgcgcaac cggatcaacg gggcgacggg gctgcgactg cccgccaccc tgggtgtcaa 9720  
 ctaccgacg ccggatgcgc tcgcgcgca cctcgtcacc gcgctgtccg cagaccgcct 9780  
 ggccgggaca ttcgaggaac tcgacaggtg ggcggcgaa ctgcccacgc tggccaggga 9840  
 tgaggccacg cggcgcgaga tcaccacccg gctaagggcg atcttgagga gcctggcgga 9900  
 cgtgtccggc ggaaccggcg gcggctccgt gccggaccgg ctgagatcgg ccacggacga 9960  
 cgagcttttc caactcctcg acaacgatct cgaacttccc tgatgcctca gccggagcct 10020  
 tcgcaacttc ctggaggga acgccacatg tcgaatgaag agaagctccg ggagtacttg 10080  
 cggcggtgcgc tcgtggatct gcaccaggcg cgcgagcggc tgcacgaggc ggagtccgga 10140  
 gagcgggaac ccatcgcat cgtggcgatg ggctgccggt acccggttg ggtgcaggac 10200  
 ccggaagggc tgtgaaaact ggtcgctcc gccggcgacg ccatcggtga attcccgcct 10260  
 gatcgtggtt ggcacctcga cgagctctac gatcccgacc cggatcagcc cggaacctgc 10320  
 tacaccggc acggcggtt cctccacgac gccggcgagt tcgacgcggt attcttcgac 10380  
 atcagcccc gtgaggcgct cgcgatggac ccgcagcagc ggctgctgct ggaaatctcc 10440  
 tgggagaccg tcgaatccgc tgggatggac ccgaggtcct tgcgggggag ccgcaccggg 10500  
 gtgttcgcy gattgatgta cgagggtat gacaccggcg cccacggggc aggagaagg 10560  
 gtcgaaggct atctcggaac cggcaatgcy ggaagcgtcg cctctggtcg ggttgcgtat 10620  
 gcgttcgggt tcgagggcc agcgggtgac gtagacacg cgtgctcgtc gtcgttggtg 10680  
 gcgctgcatt tggcgtgtca gtcgttgcy caggcgagt gtgatctggc gctggccggt 10740  
 ggagtgcag tgatgtcgac gccggagagg ttcgtggagt tctcccgta gcgtggtctc 10800  
 gcaccggatg ggcggtgtaa gtcgttcgcy gcggtgcy atggaaccgg ttggggtgag 10860  
 ggtgccggtt tgggtgtgct ggagcggtc tcagacgcca ggcggaacgg gcatcggtga 10920  
 ctggcggttg ttcgtggtag cgcggtgaat caggacggtg cgtcgaacgg attgacggcc 10980  
 ccgaacgggc tggcccagga gcgggtcatt cagcaggtgc tcacgagtgc ggggtgtcg 11040  
 gcgtccgatg tggacgctgt ggaggcgcat ggaacgggta cgcggcttg tgatccgac 11100  
 gaggcgagc ctctgatagc cgcctatgga caggatcggt accgggacgg gcgctgttg 11160  
 ttggggtcgg tcaagtccaa catcggtcat acgcaggcg ctgcggcgct cgtggtgtg 11220  
 atcaagatgg tcatggcgat gcggcacggg gagctgccgc gcacgttgca cgtggacgag 11280  
 ccgaattcgc acgtggactg gtcggctggt gcggctccgac tctgaccga gaacatccgc 11340  
 tggccaggga cgggtacgcy ccgcgtgga gtgtcgtcgt tcggggtaag cggtaacca 11400  
 gcacacgtca tctcgaaca cgaccgctc gccgtgaccg agaacgagga agcagcgag 11460  
 tccccagcac ctgggatcgt gccctggcg ttgtccgggc ggtcgtcgac ggcgctgcy 11520

gcccaggccg aacggctg cgagctgtgc gagcagaccg atcccccc cgtcgatgtc 11580  
ggtttctcac tggccgccac gcgcacggct tgggagcacc gagcgggtgt gcttggtcgg 11640  
gacagcgcta cgttgcgctc cgggcttggc gttgttgcca gcggtgaacc agcggtcgat 11700  
gtcgttgagg ggagcgtcct ggacggcgag gtcgtcttcg tcttccccgg tcagggtcgg 11760  
cagtgggccc gtatggcagt cgacctgtg gacgcttcgc cgacgttcgc gcgccacatg 11820  
gacgagtgcg ccaccgcgt ggcggaggtac gtggactggt cgttggtcga cgtgtcgcg 11880  
ggagcggaga actccccacc gctggaccgg gtggacgtgc tccagccccg gtccttcgcg 11940  
gtgatggtgt cgctcgccga ggtgtggcgt tcctacgggg tgaggccggc ggccgtcgtc 12000  
ggccacagtc aaggcgaaat cgccgcggcc tgcgcagccg ggggtgctgc gctggaggat 12060  
gcggccaggc ttgtcgcatt gcgcagcaga gcgttgaaag gactttcggg gcggggtggc 12120  
atggcgtcgc tggcctgccc tgcggatgag gtcgcggcat tgttcgcggg atcgggcccgc 12180  
cgtctggaag ttgcggcgat caacggcccc cgatcggtcg tgggtgccgg cgatctggaa 12240  
gcggtggacg aactgctggc agagtgcgt gaaaaggaca tgcgtgcacg ccgtatcccc 12300  
gtcgactacg cctcgcattc agcgacgtg gaggtggttc ggagcccgg gctggcggcc 12360  
gccgcggggg tgcgacaccg ggacggccag gtgccgtggt ggtcgacggg gatcggcgac 12420  
tgggtggatc cggccaggct ggacggcgag tattggtatc ggaacctccg gcagccggtc 12480  
cggttcgaac acgcccgtgca gggcctggtc gagcggggat tcggcctgtt catcgaaatg 12540  
agtgcgcac cgggtgctgac cagggcggtc gaggaaccg gtgcggagtc ggagaccgcc 12600  
gtggccgcgg taggtacctt gcgacgtgac tcgggcggcc tccggagggt gttgcattcg 12660  
ctggccgagg cgtacgtgcg cggcgcccacc gtggactggg ccgtggcgtt cgggggcgcg 12720  
ggccgacggc tggacctgcc gacctaccgg ttccagcgcc agcggtagtg gctggacaag 12780  
ggagctgctt ccgacgaggc tcgtgcggtc tcggaccggc cggcgggctg gttctggcaa 12840  
gccgtggcgc gccaaagacct gaaaagcgtg tccgatgcc tcgatctcga cggcgacgca 12900  
ccgctgagcg caacacttcc agccctgtcc gtctggcacc gtcaggaacg agaaagggtc 12960  
ttggcagacg gttggcggtg ccgagtcgac tgggtacggg tggccccgca gccgggtccgg 13020  
agaacgcggg aaacctggct cctggtcgtt cccccggcg gcctcgagga agcgtggtc 13080  
gaacggctga cggatgcgtt gaacacgcga gggatcagca ccctgcgcct cgacgtgcca 13140  
ccggcgggca ccagtggcga actcgcaacc gaactccgc cgcagccga cggtgacccg 13200  
gtgaaggcaa tcctgtcgt caccgcgttg gacgagcgac cccacccga atgcaaggac 13260  
gtcccagcg ggattgcctt gctgctgaac ctggtcaagg cgtcgtgta agccgacctc 13320  
agaattctc tgtggaccat cagcgtggt gcggtcaagg caggccccg agatcggtc 13380  
ctgcgcccga tgcaggcgca agcatggggt ctggggcgag tagccgact cgaacacccc 13440  
gagcgtggg gtgggctgat cgacctgcg gattcgctgg acggcgacgt cctcacgagg 13500  
ctgggcgaag cgctcaccga cggcttggcg gaagaccaac tggcgattcg ccagtcgggc 13560  
gtgctggccc ggcgactggt acccgccccg gcgaatcagc ccgctggacg taagtggcg 13620  
ccccgagga gcgcgctgat caccggcgga ctcggcgcgg tgggcgcaca ggtggcgagg 13680  
tggttggccg aaatcggagc cgagcgaatc gtgctacca gtcgacggg caaccaagca 13740  
gcaggcgccg ccgagctgga agccgaactc cgggcccttg gagcgcaagt gtccatcgtg 13800  
gcttgcgacg tgaccgatcg tgccgagatg tccgactac tggccgagtt cgacgtcacc 13860  
gcggtgttcc acgcggccgg agtcggtcgg ctgctgccgt tggcgagac cgaccagaac 13920  
ggcctggccg aaatatgcgc ggcgaaagtc cgcggcgctc aggtgctgga cgaactgtgc 13980  
gacagcaccg atctcgatgc ctctgctcgt ttctcctcgg gtgcgggggt atggggcggg 14040  
ggcggtcagg gcgcttacgg cgcggcgaaac gcattcttgg acacactcgc cgaacaacgc 14100  
cgagcacgcg gtctgccggc aacctcgatc tcctggggca gttgggcccg cggcggcattg 14160  
gccgacggcg cggcgggcga acacctgcgg cgacgcggga tacgtccgat gccggcgcg 14220  
tcggccatcc tggctctgca ggaagtactt gaccaggatg agacgtgcgt gtcgatcgtc 14280  
gatgtggact gggaccgatt cgttccccacg ttccgcgcga ctgcgcgcac ccggttgttc 14340  
gacgaagtgc cggcggcgag aaaggcgatg cccgcgaatg ggccggcaga accaggcggc 14400

tcgccgttcg cccgcaatct cgcggagctg ccggaagccc aacgacgcca cgaactggtg 14460  
 gatctggtgt gcgcccaggt ggcaaccgtg ctccggcagc gcagtcgca ggaagtccag 14520  
 cccgagcggg cgttcgcgc gctcgggttc gactccctca tggcgggtga tctgcgcaat 14580  
 cgtttgacca ccgccaccgg gttgcgcctg ccgaccacaa ccgtcttcga ctaccggaat 14640  
 ccggccgcct tggccgctca cctgctcgag gagctggtgg gtgatgtcgc gtcgggtcgc 14700  
 gtgaccgctg ccagcgcgcc cgcgagtgac gaaccgatcg cgatcgtcgc gatgagctgc 14760  
 cggtttcgg gtggcgcgca ctccgaggaa gacctgtggc ggctggtcgc cgccggcacg 14820  
 gaggtgatcg gcgagttccc ctccgaccgg ggctgggatg cggaaggcct ttacgatccg 14880  
 gatgcttcca ggcttgaac gacgtatgcg cggatggcgg gattcctcta cgacgcgggt 14940  
 gagttcgatg ccgacctgtt cggcatcagc ccacgtgagg cgttggcgat ggatccgcag 15000  
 cagcggttgg tgctcgaaat cgcctgggaa gccctcgaaac gggccggaat cgatccgttg 15060  
 tccttgaagg gcagtggggt cggcacgtac atcggcgctg gaagccgtgg gtacgcgacg 15120  
 gatgtgcggc agtttcccgga ggaggcggag ggctacctgc tgacgggtac ctccggccagt 15180  
 gtgctgtcgg gtcgggtcgc gtattcggtt ggtttcgagg gtctgcgggt gacggtggat 15240  
 acggcttggt cgtcgtcgtt ggtggcgttg catctggcgt gccagtcgtt gcgttcgggc 15300  
 gagtgtgatc tggcgttggc cgggtggtgt accgtgatgt cgacgccgga gatgttcgtg 15360  
 gagttctccc gtcagcgcg tttggcgccg gatgggcgggt gcaagtcgtt cgcggagagc 15420  
 gcggacggca ccggctgggg cgaaggcgcg ggcctgttgt tgctggagcg gttgtcggac 15480  
 gccaccgga atgggcatcg ggtgttggcg gtggttcgtg ggtcagcggt gaatcaggac 15540  
 ggcgcctcga acggactggc ggcgcgcaac ggtccgtcgc agcagcgggt gatcaaccag 15600  
 gcaactcgca atgcggctct ttcggcgctc gatgtggatg cgggtggaggc acatggcacc 15660  
 gggaccaggc tgggtgatcc gatcgaggcg caggcattga tcgcaacgta tgggcaggcc 15720  
 cgggagcggg atcggccctt gtggctgggg tcggtcaagt cgaacatcg tcatacgag 15780  
 gccgcggcgg gtgttgccgg tgtgatcaag atggtgatgg ccatgcggca cgggcagctg 15840  
 cccgcctcgc tgcacgcgga tgagcccacg tcggaggctg attggtcgtc gggggcggtc 15900  
 cggctcctcg ccgaacaggt accttggccg gagtctgacc gtgttcgtcg ggtgggggtt 15960  
 tcgtcgttcg ggatcagcgg caccaacgca catgtgatcc tcgaacaagc tacgaatgcg 16020  
 ccagatagta cagcggagac ggacaaaaca gaatccggat ctactgtcga tattccggtc 16080  
 gttccctggt tgggtgcggg aaagacgacg gattccctgc ggggacaagc cgaacgagtc 16140  
 ttgtctcagg tcgagtcctc gccggagcag cgttcgctgg atgttgctta ctcgcttgct 16200  
 tctggccgag ccgcgctgga tgaacgcgct gtcgtgctgg gtgcggaccg cggtagctg 16260  
 gttgtcggac tggcggcggt ggccgcgggt caggaggctt ctgggggtgat cagcggaaact 16320  
 cgtgcttctg ctcggttcgg gttcgtgttc tcggggcagg gtggtcagtg gttggggatg 16380  
 ggcagagcgc tctactcgaa gtttcgggtg ttcgctgctg cgtttgatga ggcttgccgc 16440  
 gagttggagg cacatctggg ggaagaccgc cgggttcggg atgtggtctt cgttccgat 16500  
 gcgcagctgc tggatcagac gctgtgggcg cagtcgggtc tgttcgcgct gcaagccggc 16560  
 ctcttggggc tgctgggttc gtggggcggt cggccggatg tggtagtggg gcattcggtc 16620  
 ggggagttgg ccgccgcgtt tgcggctggc gtgttgctg tgcgggatgc ggctcggttg 16680  
 gtggccgcgc gcgcccggtt gatgcaagcc ctgccctctg acggcgcgat gttggcggtg 16740  
 gctgctggtg aagaccttgt tcggccattg ctggccggtc gggaggagtc cgtgagcgtc 16800  
 gccgcgctca atgcccccg ttcgggtggt ttgtcgggcg atcgggaggt gctggccagc 16860  
 atcgtcggcc ggtgacga gctcggagtc cggacgcggc gtttgcgggt ctccatgct 16920  
 tttcattcgc accggatgga cccgatgttg ggcgagttcg ccagatcgc cgagtctgcg 16980  
 gagttcggta agccaacgac accgcttggt tcgacgttga cgggtgagct cgacagagcc 17040  
 gcggaaatga gcacaccagg gtattgggtg cgccaggcgc gtgaaccctg ccgtttcgcc 17100  
 gacggtgtcc aggccctggc agcgcagggc ataggcacgg tcgtcgagct cggcccgagc 17160  
 ggaacgctgg cggcactggt tcgggagtggt gcgaccgagt ccgatcgggt tgggcggatt 17220  
 tcgtcgatcc cactgatgcg cagggagcgg gacgagaccg gttcgggtgat gacagccctg 17280

gcgcattctcc acaccggtt tgggtgaggtg gactggcagg cgtttt cgggtaccggc 17340  
 gctaggcagc tgcagttgcc aacgtatgcc ttccaacgac agcactactg gatcgagtcc 17400  
 agtgcgcggc cagcacgcga ccgcgcagac atcggcgagg tggcggaaca gttctggacc 17460  
 gcggttgacc aaggcgatct ggcaacgttg gtcgcccgtc tggatcttgg ggcggacgac 17520  
 gacacatgcg catcgttgag cgatgtattg ccggcggttg cctcctggcg aagcggactc 17580  
 cgcaaccgtt cgctcgtcga ttctgcccgg taccgaatca gttggcattc ctctcgggag 17640  
 gtgccggccc cgaagatttc cgggtacctg ctggttggtc tgcgcgggtg tgcggatgac 17700  
 ggattggtca cggctttgac gagttcactg gtcggaggcg gcgcggaggc cgtccggatc 17760  
 ggccgtgtccg aagaggaccc gcaccgcgag gacgtcgac agcgggtggc caatgcgctg 17820  
 acggatgccg gtcaactcgg tggcgtgctt tcgctggttg ggctcgatga atcgctgct 17880  
 ccgggattct cctgcttgcc aactggtttc gcgctgactg tgcagcttct gcgggccttg 17940  
 cggaaggccg acgtcgaggc gccttttttg gcggtgacgc gcggcggcgt cgcgttgaa 18000  
 gatgtacgcg tgtctccgga gcaggccctg gtctgggggc tgctgcgtgt cgcgggactg 18060  
 gagcaccggg agttctgggg tggcttgatc gacctgccat cggactggga cgaccgattg 18120  
 ggtgcccggg tggcggtgtg gttggcggat ggtggcgagg atcaagtgc cattcgccgt 18180  
 ggtggtgtgt tcgtgcggcg gttggaacgc gctggtgcgt cgggtgccgg gtcggtgtgg 18240  
 cgtcctcggg ggacggtgtt ggtgacgggt ggtacggcg gtttgggggc gcatgttgcc 18300  
 cgggtggttg ccggtgccgg ggtgagcac gtggtgttga ccagccgtcg aggagcggac 18360  
 gtcggggcg ctggggaatt gcgggcggag ctggaggcgc tgggtgctcg ggtgtcgatt 18420  
 gtgccctgcg acgtggctga tcgtgacgca gtggctggag tgttggcagg gatcggtggg 18480  
 gagtgccgc tgcactgcgt ggtacacgcc gccggggctg gcgaggcggg cgacgtagtg 18540  
 gagatgggtt tggcggattt tgcagcgggt ttgtcggcga aggtgcgtgg tgcggcgaat 18600  
 ctggacgagt tgctggccga ctcgagctg gatgcgtttg tgatgttctc ctcggtgtcg 18660  
 ggggtgtggg gagccggcg acagggtgcg tatgcggctg cgaacgccta cttggatgcg 18720  
 ttggccgagc agcgtcgggc gaggggattg gtcgggaccg cggttgcgtg gggaccgtgg 18780  
 gccggtgacg gcatggccgc cggcgaaacc ggcgacagc tgcaccggat gggcctggcg 18840  
 tcgatggaac cgagcgccgc gctgctggca cttcagggtg cattggaccg cgatgagacc 18900  
 tccctcgctg tggccgatgt cgattgggca cggttcgccc cagccttcac ctcggcacgt 18960  
 cgacgcccgc tgctggacac catcgacgag gcccgagccg cattggaaac caccggcgaa 19020  
 caagcgggca caggcaaacc cgttgagctg acgcaacgcc tggccggact gtcgcggaag 19080  
 gaacgcgacg atgcggtatt ggatctggtg cgggcggaga cggcggtgt gctgggacgc 19140  
 gacgatgcca cggccctggc gccatcgcg cggttccagg aactcggatt cgactccttg 19200  
 atggcggttg agctgcgcaa ccggctgaac accgccaccg ggatccagct gcccgccagc 19260  
 acgattttcg actaccccaa tgcgagtcg ctgtcgcgtc acctctgcgc cgagcttttc 19320  
 ccaacggaga ctaccgtgga ctcgccctt gccgagctcg atcgaatcga gcagcagctc 19380  
 tcgatgtca ccggcgaagc gcgggcacgg gaccgaatcg cgacacgact gcgagccctc 19440  
 cagagaagt ggaacagcgc agctgaagta ccgacggag ccgatgtcct gagcacgctc 19500  
 gattcggcga cgcacgacga gatattcgag ttcacgaca acgagctcga cctgtcctga 19560  
 gcagttcctg cggaacttca agcgcgaaa tcgggtggaa atcacaatgg ccaatgaaga 19620  
 aaagctcttc ggctatctga agaaggtaac tgcggacctg catcagaccc ggcagccct 19680  
 gctcgcgcc gagagccgga gtcaggagcc gatcgcgac gtctcggcga gctgccgact 19740  
 gcccgccggc gtcgactctc ccgaagcgt ctggcaactc gtgcgcactg gcaccgacgc 19800  
 catctcggag ttccccgcg accggggctg ggatctcggc cggttgtacg atccccagcc 19860  
 gaaccaccag ggaacgtcgt acacgcgggc cggcggtttc ctgcaggag cgggcgattt 19920  
 cgaccccgcc atgttcggga ttctgccgcg tgaggcgttg gcgatggacc cgcagcaacg 19980  
 gttgttgctg gagctgtcct gggaggccct cgaacgggcg ggcatagacc cgacatccct 20040  
 gcgcggcagc aagaccggtg tcttcggttg tgtcacgcc caggagtacg ggccgtcctt 20100  
 gcaggagatg agccgaaacg ctggggggtt tggactcacc gggcggtagg tgagtgtggc 20160

gtcgggtcgg	gttgcgatt	cgtttggtt	tgagggctct	gcggtgacgg	tggtacggc	20220
gtgttcgctg	tcgttggtgg	ccctgcattt	ggcgtgtcag	tcgttgcggt	ccggcgaatg	20280
cgatctcgcg	ctggccggcg	gtgtgacggt	gatggcgaca	ccggcgacgt	tcgtggagtt	20340
ctcccgtcag	cgtggtttgg	ctccggacgg	gcggtgcaag	tcgttcgctg	ctgccgcgga	20400
tggcaccggg	tggggtgagg	gtgccgggtc	ggtgttgctg	gagcggttgt	cggatgcgcg	20460
gcggaatggg	cacgaggttc	tggcgggtgt	gcggggtagc	gcggtgaacc	aggacggcgc	20520
gtcgaatggg	ttgactgcgc	cgaatggtcc	gtcgcagcag	cgggtgatca	cccaggcgtt	20580
ggcgagtgcg	gggctgtcgg	tttccgatgt	ggatgcggtc	gaggcacatg	ggaccgggac	20640
cacgttggtg	gatccgatcg	aggcacaggc	cctgatcgcc	acgtacgggc	agggccggga	20700
gaaggatcgg	ccgttggtgt	tggggtcggg	caagtccaac	atcggtcaca	cgcaggcggc	20760
cgtgtgctgt	gccggcgta	tcaagatggt	cttgccgatg	cggcacgggc	agctgccgcg	20820
cacgttgcat	gtggatgagc	ccacgtcggc	ggtggactgg	tcggcgggtt	cggtcgggct	20880
tctcacggag	aacacgccct	ggccggacag	tggctcgtct	tgccgggtgg	gggtgtcgtc	20940
gttcggggtc	agcggcacca	acgcacatgt	gattctcgaa	cagtctccag	tcgagcaggg	21000
cgaaccggcc	gggcccgtcg	aaggcgagcg	ggaaccggat	gtagccgtcc	ccgtggtgcc	21060
ttgggtgctg	tcgggtgaaga	caccggaggc	tgcgcgggcg	caggccgaac	gggtgcattc	21120
gcatactcag	gaccggccgg	ggctgtcgcc	ggtggatgtg	gcgtattcgc	taggaatgac	21180
acgcgcggcg	ctggatgaac	gcgcagtggt	gttgggctcg	gaccgtgccg	cgctcctgac	21240
cgggttgagg	gcattcgccg	acggctgcga	tgcgcccgaa	gtggtttcgg	ggtctgtggg	21300
gcttggtggc	cgcgtcgggt	tcgtgttctc	gggtcagggt	ggtcagtggc	cggggatggg	21360
ccgggggctc	tactcgggtg	ttccggtgtt	cgccgacgcg	ttcgacgagg	cttgccgcgga	21420
gttggtatgca	cacctgggccc	aggaactgcg	ggttcgggat	gtggtgttcg	gttcgcaagc	21480
gtggttgctg	gatccgacgg	tgtgggcgca	gtcgggtttg	ttcgcgttgc	agattggctt	21540
gctgcggctg	ctgggttcgt	gggtgttcg	gccggatgtg	gtgttggggc	actcgggtgg	21600
tgagctggct	gcggtgcacg	cggctggtgt	gttgctcgtt	tcggaggccg	cgcggttggt	21660
ggcgggtcgc	gcccggttga	tgcaggcgtt	gccttctggt	ggtgccatgc	tcgcggtcgc	21720
tacgggtgag	tttcaggctg	atcctctgct	ggatgggggt	cgggaccgga	tcggtatcgc	21780
ggcgggtgaat	ggcccggaat	cggttgtgct	ctctggtgac	cgcgagctgc	tcaccgagat	21840
cgtgatcgg	ttgcacgac	aggggtgccg	gaccgggtgg	ttgcgggtgt	cgcatgcttt	21900
ccattcgccc	catatggagc	cgtatgctgga	ggagttcgcc	cagatctccc	gaggccgcga	21960
atatcacgca	ccggaactgc	cgtatcatctc	gaccctgatc	ggtgagctgg	acggtggtcg	22020
agtgatgggc	actcccaggt	actgggtgcg	tcaggtgcgt	gagcccgctc	gtttcgccga	22080
gggtgtccag	gcgcttgctg	gtcagggtgt	cggcacgatt	gtcgaattgg	gtccggacgg	22140
ggcgttgctg	acgttggtcg	aggagtgtgt	ggcggaatcc	gggcgggtgg	ccgggatccc	22200
gctgatgcgc	aaggaccgcg	acgaggcgcg	aaccgtgctg	gcagctttgg	cgagatcca	22260
caccctggtg	ggtgaggtgg	actggcggtc	gtttttcgcc	ggtaccgggg	cgaagcaagt	22320
cgacctgccc	acctacgcct	tccagcggca	gcggtactgg	ctggcatcca	ccgggcgtgc	22380
gggtgacgtg	accgccgcgc	gattggccga	ggcggaccat	ccgctgctcg	gtgcgggtgt	22440
tgcgttgcca	gacggcgaag	gtgtggtgct	gaccggtcgg	ttgacagcgg	gttcgcatcc	22500
gtggttgctc	gatcaccggg	tgtgggcca	aatcgtcgtc	ccggcacccg	cgtatcgtcg	22560
gctggtgtgg	cacgtcggcg	agcgcctcgg	ttgtggccgg	gtggaagaac	tggctttgga	22620
agcgcctcgt	atcctgcggg	atcatggagc	ggtccagggt	cagggtgctgg	tgggaccgcc	22680
cggggaatcc	ggagcccggg	cgggtggcgt	ctactcctgt	cctggcgagg	cgtatcgaacc	22740
cgagtggaa	aagcacgcga	cgggcgtgct	tctcccaccc	gtggccgcgc	agaaccatga	22800
gctgaccgca	tggccccggg	agaatgcgac	cgaaatcgat	gcagacgggg	tctacgcatt	22860
ccttgaagg	cacggtttcg	cgtacggacc	ggcctttaga	tgtctgcgcg	gtgcctggcg	22920
acgaggcggg	gaggtgttcg	ccgaagtcgc	attgccggat	gacatgcagg	cgggggtcga	22980
tcgattcggc	gtccaccccg	cgttgctgga	cgcggttctg	catgccgcgc	cagccgagac	23040

gtcggtggtc cagagcg cgcggtggtc gttctcgtgg cgtggg aacttcgcgc 23100  
 cactgaaagc gcggtggtgc gggcgcgcct ctcgttgact tcggatgacg aactgtcgtt 23160  
 ggtcgcagtg gacccggtg gccgattcgt ggcacggtt gattcgttg tgacccgacc 23220  
 gatctcccgg cagcaggtga ggtctggcgc gatcggatg tgctgttcg aggtggagt 23280  
 gcaccggaag gcgttggttg gaacaaccgc cggcgacgac cttgccatcg tcggtgacgg 23340  
 tcccagtttg ccggaatcgg tgcgcgcaac cgcacggttc gcgaccctgg atgagttccg 23400  
 tgcggccgtg gactcggacg ttcctgcccc gggttcgggtg ttggtcgcag ctatgtcggc 23460  
 cgaagaggtc gaggttgat ccttgccgtc gcgcgcccga gagtcgacct ccgatctgct 23520  
 ggctctcgtg cagtcgtggc ttgoggacga gcggttcgcc gaatcccgac tcgtggctgt 23580  
 cacgcgtgca gcggtgtcgg ccgaactcga ttcggacgtc gcggacctgg tgggtgcgtc 23640  
 gtcgtggggg ttgttgagtt cagcccagtc ggagaaccgg ggtcgttcg tgctgggtga 23700  
 cgtggacggc acacctgagt cgtggcaggc gttgccggcc gccgtgcgag caggagaacc 23760  
 gcagctggca cttcggcgcg gcgtggcgct ggtgcctcgg ttggcgcgac tcacggtgcg 23820  
 cgaggagggc tctccccgc aactcgacac ggacgggacc gtcctcatca cgggtggcac 23880  
 cggtgcgttg gggggagtgg ttgccgtca cctgggtggag gagcacggga ttcggcgtt 23940  
 ggtgttgga gcccgcgctg gctggaatgc gcctggagtc cacgagttgg tggatgagct 24000  
 ggcgcgcgcg ggcgcgctg ttgaggtggg ggcttgcat gtggctgacc gcaccgatct 24060  
 ggagcacgtg ctggccgcca ttcgggtcga ctggccgctg cgggggatcg tgcataccgc 24120  
 tggggtgctg gccgacggag tgatcgggtc cttgtcggcg gcggatgtgg gcacggtgtt 24180  
 tgccccgaag gtgacggggg catggcatct gcacgagttg acccgcgatc tggatctgtc 24240  
 gttcttcgtt cttttctctt cttctccgg gattgcgggt gccgcagggc aggccaaacta 24300  
 cgcggcgcg aacacgttcc tggatgcatt ggcgcgttat cgcggggcg gtgggtgcc 24360  
 tgggtgtcgt ttggcgtggg gactgtgggc gcaaccgac ggtatgacga gtggcttga 24420  
 cgcggcgctg gtggagcggg tggcgcgac gggcatcgca gaactttcca cggaggatgg 24480  
 actccgcctg ttcgatgccg cgttcgcga ggacgggct tgcgtcgtt ccgctcgatt 24540  
 ggacagggcg ctgctggtcg ggaacggacg atcgacgcg attccggcg tgttgagcgc 24600  
 gttggttctt gttcgcggcg gtgtggcgag gaaaacagcc aattctcagg ccgcggatga 24660  
 ggacgcactg ttgggtttgg tgcgggagca cgtttcggcc gtgctgggtt attcgggtgc 24720  
 ggtcgaggtt gggggcgacc gtgctttccg tgatctgggt tttgattcgt tgtctggcgt 24780  
 ggagttgcgg aaccgccttg ccgggtgct ggggtgcgg ttgccggcga ctgcggtgtt 24840  
 cgactatccg acgcgcggcg cgtggtgcg tttcctgcat caggaaactg caggcgaggt 24900  
 cgcgtccacg tcgacgccg tgaccagggc agcgagtgc gaagaggatc ttgttgcat 24960  
 tgtcgggatg ggtgtcgtt ttcgggtg ggtgtcgtc ccggaggagc tttggcggt 25020  
 ggtggccggc ggcgtggat cgggtggtg gttccagac gatcgggtt gggatctcgc 25080  
 ggcgttgatc gatcctgat ccgatcgtt cgggacctg tatgtgtgt agggcggtt 25140  
 tctgcgggac gcggcgagc tcgatgctga catgttcggc atcagcccg gtgaggcgtt 25200  
 ggcgatgat ccgcagcagc ggttgctgct ggaggtcgcc tgggaaacct tggagcgggc 25260  
 tgggatcgat ccgttctcgt tgcacggcag ccggaccggt gtgttcggcg gcttgatgta 25320  
 ccacgactat ggggcccgat tcattaccag agcaccggag ggcttcgaag ggcacctcg 25380  
 gacgggcaat gcggggagcg tgctgtcggg tcgggttgcg tttcgtttg gtttcgagg 25440  
 tctgcgggtg acggtggata cggcgtgtt gtcgtcgtt gtggcgttac acctggcggg 25500  
 tcaagcactg cgggcccgtg agtgcaatt cgccttgcg ggtggcgtc cgtgatgtc 25560  
 gacgcgacg acgttcgtg agttctccg tcaacgggtt cttggtccgg atggcggtg 25620  
 caagtcttc gcggcgccg cggatggcac cgggtggggc gagggtgcg gtctggtgtt 25680  
 gctggagcgg ttgtcggat ccggcgcaa tgggcacgag gttctggcg tgggtcgggg 25740  
 tagcgcggtg aaccaggac gcgcgtcgaa tggcttgact gcgcaaatg gtcggtcaca 25800  
 gcaaagggtg atcaccagg cactcacgag tgcggggtg tccgtgtccg acgtggatgc 25860  
 tgtggaggcg catgggacgg gcacgcggct tggatgccg atcgaggcg aggcgttgat 25920

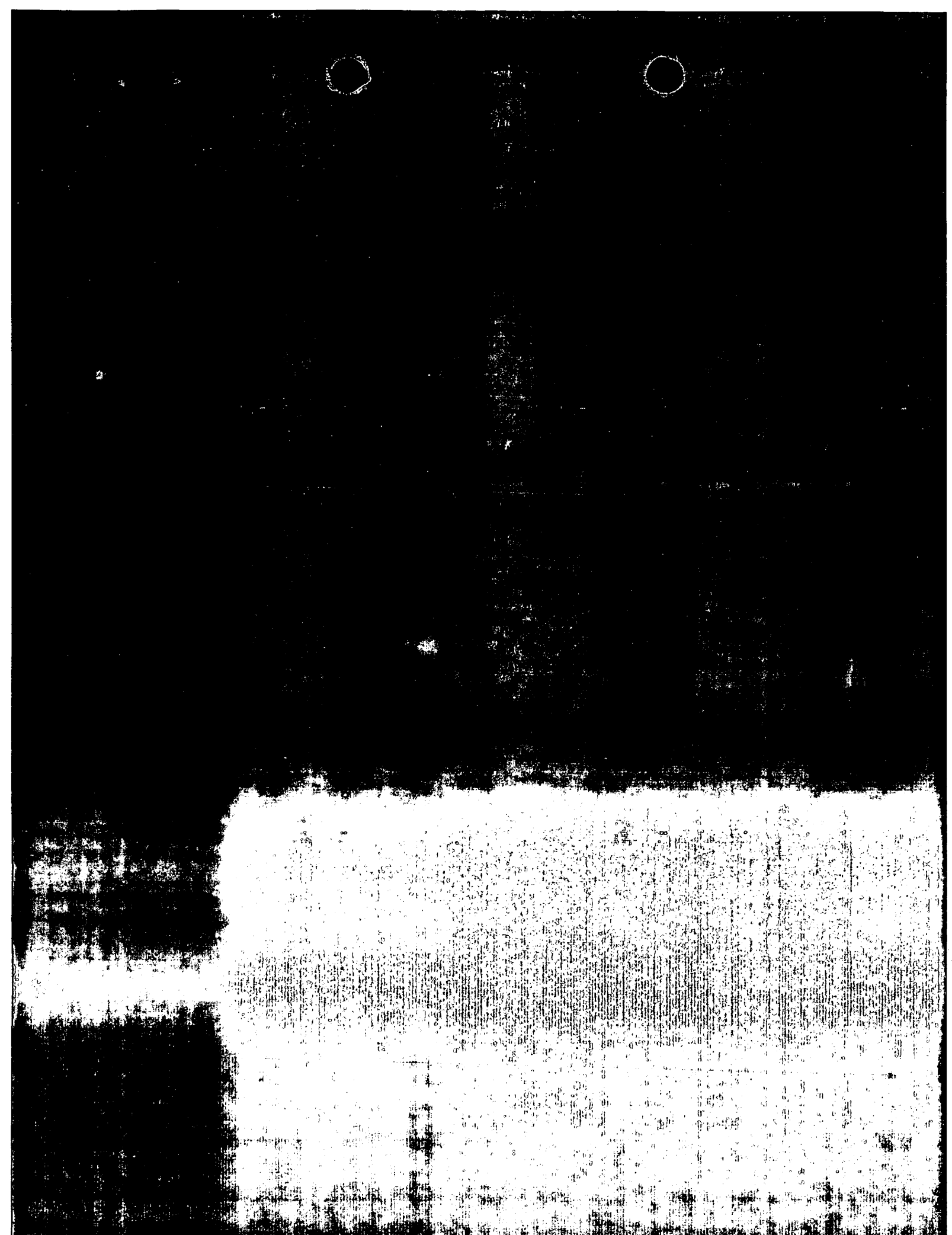
cgctacgtac	ggccgggatc	gtgatccccg	tccgcccgtt	tggctggggt	cggtgaagtc	25980
gaatattggt	cacaccagg	cggcggcg	tgctcgctggt	gtgatcaaga	tggtgatggc	26040
gatgcggcag	ggggagctgc	cgcgcacgtt	gcacgtggac	gagccctccg	cgcaggtgga	26100
ctgggtctgcg	ggcacggtcc	aactcctcac	ggagaacacg	ccctggcccc	acagcggtcg	26160
tcttcgcgg	gcgggctgt	catcggttcg	gatcagtggc	accaacgcgc	acctgatcct	26220
tgaacaacct	ccgcgagagt	cgcagcgctc	aacagagccg	gattcgggtt	ctgtccgcga	26280
ttttccggtg	gtgccgtgga	tggtgtcggg	caaaacaccc	gaagcgctat	ccgcccaggc	26340
agatgcattg	atgtcctact	tgagcaatcg	cgttgatgct	tccccgcgag	atatcggtta	26400
ttcgcttgog	gtgaccgctc	cggcgttgga	ccaccgcgct	gtcgtgctgg	gtgcggatcg	26460
tgccgcgttg	ctgccgggct	tgaagcgct	ggccgttagt	aatgacgctg	ccgaggtgat	26520
caccggcact	cgtgccgctg	ggccggtcgg	attcgtgttc	tccggtcaag	gtggtcagtg	26580
gcccgggatg	ggaagcgggc	tccactcggc	gtttccggtg	ttcgccgacg	cgtttgacga	26640
agcctgctgc	gagctggatg	cgcactctcg	gcagatggcc	cggctacgag	atgtgttgc	26700
cggttcggat	acgcaacttc	tggaccagac	cttgtgggcg	cagccgggcc	tggtcgcgtt	26760
gcaagtccga	ctctgggagt	tggtgggttc	gtggggtgct	cggcccgcctg	tggtgctggg	26820
ccactcggtc	ggtgagctgg	cggcggcggt	cgcggctgga	gtgttgctgt	tgccggatgc	26880
ggctcggctg	gtggcgggcc	gtgcccggtt	gatgcaagcc	ctgccaactg	gcggtgccat	26940
gctcgtctgc	gctgctggag	aggagcagct	gcgcccggtg	ctggccgact	gcggtgatcg	27000
tgtggggatc	gccgcggtca	acgctcccgg	gtcgggtggtg	ctctccggtg	atccggatgt	27060
gctcgatgac	attgccggtc	ggctggacgg	gcaagggatc	cggccaggtt	ggttgccgggt	27120
ttcgcatgog	tttcattcgc	atccgatgga	tccgatgctg	gcggagttca	ccgaaatcgc	27180
ccggagcgtg	gactaccggt	cgtcagggct	gccgatcgtg	tcgacgttga	cgggtgagct	27240
cgatgaggtc	ggcatgccgg	ctacgccgga	gtattgggtg	cgccaggtgc	gagaacccgt	27300
ccgcttcgcc	gacggtgttg	ctgcgctcgc	ggctcacggt	gtgagcaccg	tcgtcgaggt	27360
cggtcgggat	ggggtgttgt	cggcgctggt	gcaggagtgc	gcggccggat	ccgatcaggg	27420
cggacgggtg	gccgcggttc	cgtcatgcg	cagcaatcgc	gacgaggcgc	acacggtgac	27480
aacggcattg	gcgcagatcc	atgtgcgtgg	tgctgaggtg	gactggcggt	cgtttttcgc	27540
cggtagccgg	gcaaagcagg	tcgagctgcc	cacgtatgcc	ttccaacgac	agcgggtactg	27600
gcttgactca	ccatccgaac	cggtcgggca	atccgccgat	cccgcgcgcc	agtcgggctt	27660
ctgggaactc	gtcgagcagg	aagatgtcag	cgcgctcagc	gccgctctgc	acattaccgg	27720
cgatcacgac	gtgcaggcgt	ccctggaatc	ggtgggttccg	gtcctctcct	cctggcatcg	27780
ccggatccgc	aacgaatccc	tggtgcacca	gtggcggtac	cggatttctt	ggcatgagcg	27840
ggcagatttg	ccagacccct	cgttgtcggg	gacatggctc	gtcgtcgtgc	cggaggggtg	27900
gtcggcgagt	cggcaagttc	tgcgtttcaa	cgagatgttc	gaggaacggg	gttgcccggc	27960
agttctgttc	gagctcgccg	ggcacgacga	ggaagccctg	gcgcaacgat	tccgctcgtt	28020
gcctgttgog	tcagggggaa	taagcggcgt	gttgtccttg	ctggcgctgg	atgaatcgcc	28080
gtcctcgccg	aacgctgctt	tgccgaatgg	cgcgctgaac	tcgttggtac	tgctgcgagc	28140
tctgccccgc	gcggatgtgt	cggcgccatt	gtgggttgccg	acgtgtgggtg	gtgtcgcgggt	28200
cggggatgtg	ccggtgaacc	cggggcaggc	gctggtgtgg	ggactgggtc	gcgtcgtcgg	28260
tctggagcat	ccggcctggt	ggggtggcct	ggtcgacgtg	ccgtgcttgc	tcgatgagga	28320
cgtcagagaa	cgtttgtcgg	tcgtgttgcc	aggtccttggc	gaggacgaga	tcgcggtacg	28380
tcccgggtgt	gtgttcgtgc	ggcgggttga	acgcgctggt	gcggcgctcg	gtgcggggtc	28440
ggtgtggcgt	cctcggggga	cgggtgttgt	gacgggtggt	acgggcggtt	tgggggcgca	28500
tggtgccccg	tggttgccgg	gtgccggggc	tgagcatgtg	gtgttgacca	gccgtcgagg	28560
cgcggcggtc	ccgggcgctg	gagatttgcg	ggcggagctg	gaggcgctgg	gcgctcgggt	28620
ttcgatcacg	gcctgcgacg	tggccgatcg	tgacgctttg	gccgaagtgt	tggcgaccat	28680
tccgatgat	tgcccgctga	ccgcggtgat	gcctgcggcg	ggggtcgttg	aagtcggcga	28740
cgtggcgctc	atgtgtttga	ccgacttcgt	tggggtgctg	tcggcggaag	caggtggtgc	28800

ggccaatctc gatgagtc tcgccgatgt cgagctggat gccttcgc tgttctcacc 28860  
 cgtctcgggt gtgtggggtg ctggcgggca gggcgcttat gcggcggcga atgcctactt 28920  
 ggatgcgttg gcgcagcagc gtcgggcaag ggggttggtg gggactgcgg ttgctggtgg 28980  
 cccgtggggc ggtgacggaa tggccgcagg tgaaggcggt gcacagctgc gccggggccg 29040  
 cctggtgccca atggctgcgg atcgggcgtt gctggcactt cagggcgcat tggatcgtga 29100  
 cgagacatcc ctggtcgtgg ccgatatggc gtgggagagg ttccggcccg tgttcgccat 29160  
 gtcccgtcgg cgtccgctgc tcgacgagct gcccgaaagca cagcaggcgt tggcggtatgc 29220  
 ggagaacacc actgatgctg cggactcggc cgtcccgcga ccgcggctcg cgggcatggc 29280  
 agccgcgaa cgccgcgcg cgatgctgga cctggtgctg gcggaggcct cgattgtgtt 29340  
 gggacacaac ggttctgacc cagttggtcc cgaccgggcg ttccaggagc tcggatttga 29400  
 ttccgtgatg gccgtcgaaac tgcgcaacag gttgggagag gcaacaggat tgagtctgcc 29460  
 ggccacgttg atcttcgatt atccgagccc atccgcgctg gctgagcagc tggtcggcga 29520  
 gctggtggga gcgcagccc cgaccaccgt cgtggccggg gccgatccag tggatgatcc 29580  
 ggttgctcgtg gtcgcgatgg gatgccggtg tccgggagac gtctgctcgc ccgaggagct 29640  
 gtggcagctg gtttctgcgg gacgtgatgc ggtatcgacg ttcccgcgtg atcgggggtt 29700  
 ggactgcaac acgttggttcg acccggtacc ggatcgggca ggcagtacct atgtgcgaga 29760  
 aggtgccttc ctgaccggtg ctgatcggtt cgacgcgcgg ttcttcggca tcagccctcg 29820  
 cgaggcgcg gcaatggatc cgcagcagag gttggtgctc gaagtggcgt gggaggtttt 29880  
 cgaacgagca ggaatcgctc cgctgtcgtt gcggggtagc aggaccggtg tgttcgcggg 29940  
 gaccaatggg caggaccacg gtgcgaaagt ggctgcgcgc ccggaggcgg cgggtcacct 30000  
 cctgaccgga aacgcgcga gtgtcctggc cggccgggtt tcctacacgt tcggccttga 30060  
 ggggcctgcg gtggcggtgg ataccgcgtg ttgctcgtcg ttggtggcgt tgcatttggc 30120  
 gtgccagtcg ctgcgttcgg gtgagtgtga tatggcgttg gcagggtgtg tgacgggtgat 30180  
 gtgcacaccc ctggctttcc tcgagttctc tcgtcagcgc ggtttggcgc cagatggtcg 30240  
 gtgcaagtcg tttgcggccg ctgcggatgg caccgggtgg ggtgaggggt ccggcctggt 30300  
 gttgctggag cggttgctcg atgctcgtcg gaatggcac ccgggtgttg ccgtggttcg 30360  
 cgggtctgcg gtgaatcagg atggtgcgtc gaatggcctg actgcgcga atggtccgtc 30420  
 gcagcagcgg gtgattcggc aggcctcgc gaatgcgggg ctgtcggcgt ccgatgtgga 30480  
 tgtcgtggag gcgcacggga ccggtaccgg gctcggggat ccgatcgagg cgcaggcgct 30540  
 gatcgcgaca tatgggcagg agcgggatcc tgagcgggcc ctgtggctgg ggtcgatcaa 30600  
 gtccaacatc ggccacacgc aggcggcggc cgggtgtggc ggggtcatca agatggtgca 30660  
 ggccatgcgg cacggggagt tgcctgcgac gttgcacgtg gacaagccca ctccacaggt 30720  
 ggactggtct gccggggccg ttccggtcct caccgggaac acgccctggc ccgagagcgg 30780  
 ccgtcctcgt cgagcggggg tgtcgtcgtt cgggatcagc ggcaccaacg cacacctcat 30840  
 cctcgaacaa ccaccgtcg aaccagcgga gatcgaccaa tcggatcggc gggtcactgc 30900  
 gcatccagcg gtgatccgt ggatgttgtc ggctaggagt ctgcagcgc tgcaggccca 30960  
 agcggctgcg ctgcaggccc ggctggaccg gggctcctggc gcttctccgc tggatttggg 31020  
 gtattcactc gcgaccactc gttctgtgct ggacgaacgc gccgtcgtgt ggggtgccga 31080  
 tcgggaggca ctgctgtcca ggctggcagc gctcgccgat ggccggacgg ccgggggggt 31140  
 gataacgggc tctgcgaatt ccggtggccg catcggtatc gtttttccg gtcagggcag 31200  
 tcagtggctg gggatgggaa aggcgttgtg cgcggcttcc ccggcgttcg cggacgcctt 31260  
 cgagggaagc tgcgacgcgc taagcgcaca cctgggcgcg gacgttcggg gtgtgctgtt 31320  
 cggtgctgat gagcagatgc tcgaccggac gctgtgggag cagtccggga tcttcgcgggt 31380  
 tcaagtcggc ctccctggat tgcgtgaggtc gtggggcgctg ccggccggccg cgggtgctggg 31440  
 gcaactcggc ggagagttgg ctgcggcgca cgcggctggt gtgttgtcct tgccggacgc 31500  
 tgcacggttg gttgcggctc gggcccaact gatgcaggca ttgccaccg gcggcgcaat 31560  
 gctcgcggtc gccaccagcg aggcggcggt cggaccgctg ctttccgggg tgtgcgatcg 31620  
 ggtcagcatc gctgcgatca acggccccga gtcggtagtg ctctccggcg accgcgatgt 31680



gctcgtggag	ctcgcaggcg	aattcgaatgc	ccgagggcctt	aggaccaa	ggttgcgggt	31740
ctcccatgct	ttccactcgc	accggatgga	accgattctg	gacgagtacg	cggaaccgc	31800
caggtgcgtc	gagttcgggtg	aaccgggtggt	gccgatcgtc	tccgccgcga	ccggtgcgct	31860
ggacaccacc	ggactgatgt	gcgcggccga	ctactggacg	cgccaagtgc	gtgatcctgt	31920
ccgcttcgga	gacggtgtcc	gggcgctcgt	cggccaaggc	gtggacacga	tcgtcgagtt	31980
cggcccggac	ggggcgttgt	cggccctggt	cgagcagtg	ttggccgggt	ccgaccaggc	32040
tgggaggggtg	gcggcgatcc	cgctgatgcg	cagggaccgc	gatgaggtcg	agaccgcggt	32100
ggcggccctg	gcgcacgtgc	acgtccgcgg	tgggtcggtg	gactggtcgg	cttgcttcgc	32160
cggcaccggc	gcccgcaccg	tcgagttgcc	cacctacgcc	ttccaacgcc	agcggtaactg	32220
gctggccggg	caagcggacg	ggcgcggcgg	cgatgtggtt	gccgaccggg	tcgacgcgcg	32280
cttctgggag	ttggtcgagc	gcgcgatcc	ggaaccggtg	gtggatgaac	tctgcatcga	32340
ccgggaccag	cccttcggg	aggtgctgcc	cgttctggt	tcctggcgcg	agaaacaacg	32400
ccaggaggcc	ctcgcggatt	cctggcgcta	ccagggtgcg	tggaggtccg	tcgaggtgcc	32460
gtccgcagcc	gccctccggg	gcgtgtggt	ggtggtgctt	ccagctgacg	tgccccgaga	32520
tcaaccggcg	gtcgtcatcg	acgcgctgat	cgcgcgcggc	gccgaggtcg	cggtcctgga	32580
attgaccgag	caggacctcc	aacgcagtg	gcttgtggac	aaggtgcgcg	ccgtcattgc	32640
ggaccgcacc	gaggtgacgg	gtgtgttgtc	tctgttggcg	atggacggca	tgccctgcgc	32700
ggcgcacccg	cacctgtccc	gtggtgtcgc	cgctaccgtg	atcctgacgc	aggtgttggg	32760
cgatgcgggt	gtttccgccc	cgctgtggt	ggccacgacc	ggtggcgtcg	aggccgggac	32820
cgaggacggt	ccggccgac	cggaccacgg	cttgatctgg	gggctcggca	gggtcgtcgg	32880
ccttgaacat	ccgcagtggt	ggggtggcct	gatcgacctt	ccggagacac	tggacgagac	32940
gtcccgaac	gggttggtgg	ccgcactcgc	cgggacggcg	gccgaagatc	agctcgccgt	33000
gcgttcaccc	gggttggtcg	ttcgcagagt	ggtgcgcgca	gcgcggaacc	cccggtcaga	33060
gacatggcgt	agccggggaa	cggtcctcat	cacgggcgga	acaggcgcgc	tcggtgccga	33120
ggtcgcacga	tggctggccc	ggcggggagc	tgagcacctg	gtgttgatca	gtcgccgcgg	33180
cccggaagct	cccggcgcag	cggacctagg	ggccgagctg	actgaactcg	gcgtgaaagt	33240
cacagtcttg	gcctgcgatg	tgacggaccg	cgacgagctg	gcggcggtgc	tggcggccgt	33300
tcccacggag	tatccgctgt	cggcggtcgt	gcacaccgcc	ggcgctcggga	cgctgcgaa	33360
cctggccgag	acgaccttgg	cgcagttcgc	cgacgtgttg	tcggccaagg	tcgtcggcgc	33420
ggcgaacctg	gaccggctgc	ttggcgggca	accgttgga	gccttcgtgc	tggtctcctc	33480
gatctcggga	gtttggggag	ccggcgccca	aggagcctat	tcggccgcca	atgcgtatct	33540
cgatgccctt	gccgagcgcc	gacgggcttg	cggcgggccg	gcgacgtgca	tcgcctgggg	33600
tccgtgggcg	ggtgcgggca	tggccgttca	ggaaggtaac	gaggcgcac	tcgcgcaag	33660
gggcctggtg	ccgatggaac	cgcagtcggc	cctcttcgcg	ctgcaacagg	ccctgtccca	33720
acgagaaacc	gccatcacccg	tcgcagatgt	ggactgggag	cgattcgcgc	cctctttcac	33780
cgcggcccg	ccgcgaccac	tggttgaaga	gatcgtggat	ctacggccccg	acaccgagac	33840
cgaggagaag	cacggtgccg	gcgagctggg	gcagcagctg	gccgcactgc	cgcccgctga	33900
gcgcggacac	ctgctgctgg	aggtggtgct	ggcggaacc	gccagcacc	tggggcacga	33960
ttcggcgag	gctgtgcaac	ccgatcggac	cttcgccgaa	ctgggcttcg	attcgtgac	34020
cgcggtagag	ctgcgcaaca	ggttgaacgc	ggtgaccggg	cttcgcctgc	cgccgacgct	34080
ggttttcgac	caccgcacgc	cgctggcggt	gtccgaacag	ttggttcggg	ccctggtcgc	34140
ggagccggac	aacggcatcg	aatcgctgct	cgcgagctc	gacaggctgg	ataccacgtt	34200
ggcgcaagg	ccttcgatcc	cactggaaga	ccaggccaag	gtggcggagc	gcttgcaacgc	34260
actcctcgcc	aagtgggacg	gggcgcgtga	cggcacggcc	agagcgacgt	caccccaatc	34320
gctgacggcg	gccacggacg	acgaaatctt	cgacctcatc	gaccggaagt	tcggcgctg	34380
accgcccttt	cctgcctca	gctccctga	ttactggaac	ggtgtatttc	gatggccaat	34440
gaagaaaagc	tcgcgagta	cctcaagcgt	gtcgtcgtcg	aactggaaga	ggcgcacgaa	34500
cgctgcacg	agttggagcg	ccaggagcac	gaccccatcg	cgatcgtgtc	gatgggatgt	34560

cgttatcccg gtggcggtctc cactccggag gagctgtggc gactgggtcgt cgacggagga 34620  
 gacgcgatcg cgaacttcc aagaccgt ggctggaatc tggacgac gttcgatcct 34680  
 gatccggggcc gagccgggac ctccctacgtc cgcgaggggtg gtttccctgct cggggtcgctg 34740  
 gacttcgatg ccgggctctt cgggatcagt ccgcgcgagg cacaggcgat ggacccgcaa 34800  
 cagcgggttg tgctggagat ctctgtgggag gtgttcgagc gcgcggcat tgacccgttt 34860  
 tctttgcggg gtaccaagac cgggtgtgtc gcgggacctga tctaccacga ctacgcgtcg 34920  
 cggtttcgca agacccccgc ggagttcgag ggttacttcg ccaccggcaa cgcgggcagc 34980  
 gtcgcacccg gccgggtggc ttacaccttc gggtagagg gcccggcggg caccgtggac 35040  
 accgcctgct cgtcgctccct ggtggcgctg cacctggcct gccagtcctt gcggctgggc 35100  
 gaatgcgacc tggccctggc cgggtggcatt tcgggtgatg ccacgcggg agccttcgtc 35160  
 gagttcagcc ggcaacgcgc actcgctcgt gatggcgggt gcaagccctt cgcggatgcc 35220  
 gccgacggca ccgggtggg cgaggggcgcc ggaatgctgc tgctggaacg gctgtcggac 35280  
 gcacgacgaa acggccaccc ggtgctggcg gcgggtggcg gttccgcgat caaccaggac 35340  
 gggacgtcca acggcctgac cgcgcccagc ggtcccgac agcagcgagt gatccgcaa 35400  
 gccctggcga acgcggggtt gtgcgccgcc gaggtcgatg tggtcgaggc gcacggcacg 35460  
 ggcacggcct tgggcgaccc gatcgaggcg caggccctga tcgccacctt cggggcgaa 35520  
 cggtcggcg atcatccgt gctgctgggt tccctcaagt cgaacatcgg ccacaccag 35580  
 gctgccgccc gtgtggccgg ggtgatcaag tcggtcctgg ccatcaggca ccgggagatg 35640  
 ccccgagcc tgcacatcga ccagccatcg cagcacgtgg actggtcggc gggcgcggtg 35700  
 cggctgctca cggacagcgt tgactggccg gatctcggca ggccgcgcc agcaggggtg 35760  
 tctcgttcg gcatgagcgg taccaacgca cacctgatcg tcgaggaagt atccgacgag 35820  
 ccggtctcgg gcagtaccga gccgaccggg gcatttccct ggccgctgtc cggcaagacg 35880  
 gagacggcat tgcgcgagca ggctgccgag ttgctctccg tagtgaccga gcaccggag 35940  
 ccgggactgg gggacgtcgg gtactcgctg gccaccggtc gcgctgcgat ggagcaccg 36000  
 gctgtcgtgg ttgccgacga tcgggactct ttcgtcgccg gactgacggc gttggctgctg 36060  
 ggcgttccgg cagccaacgt ggtgcagggc gcggccgact gcaagggaaa ggtcgcgttc 36120  
 gtgttcccc gccagggtc gcattggcag gggatggcga gggaactgtc cgaatcctcg 36180  
 ccggtgttcc ggcggaagct ggcggaatgc gcggcgcta cggcccctta cgtggactgg 36240  
 tcgctgctcg gcgtccttcg cgggtgatcc gatgcacccg cgctggatcg cgacgacgtg 36300  
 attcagctcg cgtgttctgc catgatggtg tcgctggccg aactgtggcg ttcgtgcgga 36360  
 gtggagccc cgcgggtggt cggtcattcc cagggcgaga tcgccgccgc ccatgtggca 36420  
 ggcgctttgt ccttgactga tgcggtgcgc atcatcgctg cccgctgcga tcggtgtcg 36480  
 gcgctgaccg ggaagggagg catgctcgcg attgccttgc cggaaagcgc ggtggtgaag 36540  
 tgaatcgag gcctgccgga gctgaccgtt gcggcggtca acggacccgg ctccactgtc 36600  
 gtttccggcg aaccgtcggc tctggagcgt ctgcagaccg aactgaccgc ggaaaacgtg 36660  
 cagacccggc ggggtgggaat tgattacgcc tcgcattcgc cgcagatcgc gcagggtccag 36720  
 ggccggcttc tggaccggct gggcgaagtc gggtcggaac ctgctgagat cgctttctac 36780  
 tcgacgggtca ccggcgagcg gacggacacc ggccgactcg acgcccacta ctggtaccag 36840  
 aaccttcggc agcccgtccg cttccagcag accgtcgccc ggatggcaga tcagggttat 36900  
 cggttcttcg tcgaggtgag cccgcaccgc ctgctcaccg ccggaatcca ggaaacgtg 36960  
 gaagccgcgg acgcgggagg ggtggtggtc ggttcgctgc ggcgtggcga gggcggtcc 37020  
 cggcgctggc tgacttcgct ggccgagtgc cagggtgcgc gactgccggg gaattgggaa 37080  
 caggattacc tcaacaccgg agcccagcgc gtgccgctgc cgacctacc gttccagcgg 37140  
 cagcggtaact ggttgaggtc cgcgagtagc gacgcgggag atctcgggtt ggtgggcttg 37200  
 ctctccgcgc agcatccct gctcggggct gcgggtgacgc tggccgatgc gggcggttc 37260  
 ctgctgaccg gcaagctgtc ggtcaagacc cagccctggt tggccgacca cgtggtcggc 37320  
 gggcgatcc tgctgcccg caccgcgttc gtggaaatgc tgatacgcgc cgcggaccag 37380  
 gtcgggtgcg atctgatcga ggagttgtcc ctgacgactc cgctgggttt gcccgcgacc 37440



ctggcgggtc aagcac gtctggtgag tgtgatctgg ctctt gg tgggtgacg 40380  
 gtgatggcca cgccggggat gttcgtggag ttttcgcgtc aacggggctt ggcggcgat 40440  
 gggcgggtgca agtcgtttgc ggcggctgcg gatggcaccg gttggggaga aggcgcgggc 40500  
 ttggtgttgt tggagcggct gtcggatgcc cggcgcaacg ggcacgcggt tctggcggtc 40560  
 gtgcggggta ggcgggtgaa tcaggatggt gcgtcgaatg gtttgacggc gccgaatggg 40620  
 ccctcgcagc agcgggtgat cacgcaggcg ttggcgagt ctggtttgtc ggtgtctgat 40680  
 gtggacgccg tggaggcgca tgggactgga accaggcttg gtgatccgat tgaggcgag 40740  
 gctctgattg ccacttacgg gcagggggcg gatagcgatc ggccgttgtg gttggggtcg 40800  
 gtgaagtoga atattggtca tacgcaggcg gcggcggtg tcgctggtgt gatcaagatg 40860  
 gtgatggcga tgcggcacgg gcagctgcc gcgacgttg atgtggatga acctacgtcg 40920  
 gaagtggatt ggtcggcggg ggatgtccag ctctcacgg agaacacccc ctggcccggc 40980  
 aacagccatc ctccggcggt gggcgtgtcg tcgttcggga tcagcggcac caacgcacac 41040  
 gtcctcctcg aacaagcctc gaaaacacca gacgagactg cggacaagag cgttcccgat 41100  
 tcggaatcga ccgtggacct tccagcggtc ccgttgatcg tgcgggggag aacaccggca 41160  
 gcgctcagcg ctccggcgag cgcattgttg tccattttgg gtgagcgtgg cgatatttcc 41220  
 acgctggatg cggcgttttc gttggcttcc tcccgggccg cgttgaggga gcggcggtg 41280  
 gtgctgggag cggaccgcga aacgttgttg tccgggttg aagcgtggc tccgggtcgc 41340  
 gaggcttctg ggtggtgtc gggatcccc gtctctggcg gggttgggtt cgtgttcgcc 41400  
 ggtcagggcg gacagtgtt ggggatggc cggggctct actcgtttt tccggtgttc 41460  
 gctgacgcgt ttgacgaagc atgtgccgga ctggacgcgc atctggggca ggacgtggg 41520  
 gtccgggatg tgggtgtttg ttccgacggg tccgtgttg atcggacgt gtgggccag 41580  
 tcgggtttgt tcgcttgca ggttggttg ctgagcctgc tgggttcgtg ggtgtccg 41640  
 ccgggtgtg tgcgtggcca ttcggtcggc gagttcgcg cggcggttg gcgggagt 41700  
 ttgtcgttg cggatgcggc tcggatggt gcgggtcgt cccggttgat gcaggcgtt 41760  
 ccttctggcg gtgccatgtt ggcgggtggt gctggtgagg agcagctgc gccgttgtt 41820  
 gccgatcggg ttgatggtgc gggatcgc gccgtcaacg ctctgagtc ggtggtgtc 41880  
 tccggcgatc gggaggtgct tgacgacatc gccggcgcg tggatgggca agggattcgg 41940  
 tggcggcggt tgcgggttt gcctgcgtt cattcgtatc ggatggacc gatgttcag 42000  
 gatttcgccg aaatcgcac cagcgtggac taccggcgt gcgacctacc ggtcgtgtc 42060  
 acgttgacgg gtgagctcga caccgcagg gtgatggcta cgcggagta ttgggtgct 42120  
 caggttcag agcccgctc ctccgcgcac ggcgtccgg tgcgcgcga gcaaggggtc 42180  
 gccacgatc tcgaactcgg ccctgatgc acgctgtcg ccctgattcc cgattgtcat 42240  
 tctggtgctg atcaggccat gccgattcc atgctgcgta aagaccgtac ggaaccgaa 42300  
 actgtggtc cgcgggtgg gcgggcgcac acgctggtg ttccggtcga atggtcggc 42360  
 tatttcgcg gcaccgggg acggcggtc gagttgccga cgtatgcct ccagcggcag 42420  
 cggtaactggc tggaaacatc ggattacggc gatgtgacgg gtatcggcct ggtcgcggc 42480  
 gagcatcgt tgcgtggggc cgtggttgc ctggccgat gtgatggggt ggtgctgacc 42540  
 ggccggttgt cgggtgggac gcctcgttg ctggcccagc atcgcgtgct gggcgaggc 42600  
 gtcgtcccc gcaccgccat cctggagat gccctgcac caggggcgg tctcggctgt 42660  
 gaccgggtg aagagctcac cctggaaaca ccgctggtg tccccgaac cgcggcggt 42720  
 gccggtatg gtggccctgc gggagggacc acagtttcaa ttgaaactgc ggaagaacgt 42780  
 gtgcggacga acgacgccat cgaatccag ctgctggtga acgcaccga cgaaggcgt 42840  
 cggcgaagg tgcgtgtga tccccccc gccggtgggt cgagaggtg gggttggac 42900  
 cgccacgcca cggcgaact cgtcgtcgg accaccggt gtaggcggt tctgatttg 42960  
 tcggctgagg gtgccgagtc gattgtctc gatgagttc acgtcgtct ggccgaaac 43020  
 gggttcagat acgggccgtt gttccaggg cttcaggcg catggcgtc tggtagcag 43080  
 gttctcgcg aaatcgcccc gcggccgag gccgatgcga tggcgtcggg atacctgtc 43140  
 gaccagcgt tgcgtgatgc cgcgtgcag gcgtccgcgc tcggcgacc cccggagcaa 43200

ggcgggcgct	ggctgcccgt	ctcattcacc	ggcgtcgaac	tttccgctcc	ggcagggacg	43260
atcagcaggg	tgcgggtgga	gaccaggcga	cccagcgcga	tatcggtagc	cgtgatggat	43320
gagagtgggc	ggttgctcgc	ctcgatcgat	tctctcaggc	tacgaagcgt	gtcgtcggga	43380
cagctggcga	atcgggacgc	tgtccgcgac	gcgctgttcg	aggtgacctg	ggagccggtg	43440
gcgacgcagt	cgacggaacc	gggtcgctgg	gccctgcttg	gtgatactgc	ctgcggtaaa	43500
gacgatctca	tcaaactcgc	aacggattcc	gccgaccgct	gcgcggatct	ggcggcgcta	43560
gccgagaaac	ttgattccag	cgcgctgggt	cctgatgtcg	tgggtctactg	cgccggagaa	43620
caggcggatc	ccggcaccgg	cgagccgcga	cttgccggaga	cccagcagac	gttggtctctg	43680
ctccaagcgt	ggttggctga	gccgcgggtg	gccgaggcac	gtctggtggt	ggtgacgtgt	43740
gcagcgggtg	cgacggctcc	gagtgaaggc	gcacgagagc	tggcacatgc	gccgttgttg	43800
gggttggtgc	gtgccgcgca	ggtggagaa	ccggggcagt	ttgtgctggc	ggacgtcgac	43860
ggaaccgccg	aatcgtggcg	tgcgttgccg	agtgcgttgg	gctcgatgga	accgcagttg	43920
gccctgcgga	agggcgcggt	gcgagcgccc	cgcttggtct	cggtcgccgg	gcagatcgac	43980
gtgcccgcgg	ttgtggcgga	tcccgaaccg	accgtgctga	tttcggggcg	cacgggcctg	44040
ttggggggcg	cggttgcccc	ccacctgggt	accgaacgcg	gtgtccgcgc	attggtgttg	44100
acggggccgt	gtggctggga	tgtccctgga	atcaccgagt	tgggtgggtga	gctgaacggc	44160
ctcgggtgcc	tggctgacgt	ggtggcgctg	gacgtcgccg	atcgtgctga	tctggagtcg	44220
ttgctggcg	cggtcccggc	ggaatttccg	ttgtgcggcg	tgggtgcatg	cgcgggggcg	44280
ctggccgacg	gggtgatcga	gtcgttgtca	ccggacgacg	tgggagcggt	gttcggcccc	44340
aaggcggcg	gggcgtggaa	tctgcacgag	ctgactcgtg	atacggacct	gtcgtttctc	44400
gcgttggtct	cctcgctttc	cggtgttgcc	ggcgctcctg	gtcagggcaa	ttatgcggcg	44460
gcgaacgcgt	tcttggaacg	attggcgcat	taccggcggt	cacagggact	gcctgcgggt	44520
tgcgtggcct	ggggccctgt	ggagcagccg	agcgggatga	cggagacgct	cagcgagggt	44580
gaccggagca	ggatcgcgcg	cgccaacccg	ccgttgtcca	ccaaggaggg	attgcggctg	44640
ttcgatgccg	ggctggcgct	ggaccgggca	gcggtagtgc	cggcgaagtt	ggacaggact	44700
ttcctggccg	agcaggcgcg	gtcgggctcg	ctgcccgcac	tgttgacggc	actggtacct	44760
cccatccgtc	gtaataggcg	ggctagcgga	accgagctcg	cggacgaggg	cacctgctc	44820
ggggtggtgc	gggagcatgc	cgcgcccggt	ctggggattt	cgagcgccgc	tgacgtcggg	44880
gtcagagcgc	ctttccggga	tctgggtttt	gattcggtgt	ctggtgtgga	gttgcggaac	44940
cgcttgccg	gggtgctggg	ggtgcgggtg	ccggcgactg	cggtgttcga	ctatccgacg	45000
ccgagggcgc	tggcccgggt	cctgcaccag	gaactggcag	acgagatcgc	tacgacgcca	45060
gcgcgggtga	cgacgaccag	ggcaccgggt	gccgaagacg	atctcgctgc	gatagtcggg	45120
atgggatgcc	gttttcccg	tcaggtgtcc	tgcgcggagg	agctctggcg	tttgggtggc	45180
gggggcgtgg	atgcggctgc	ggacttccca	gccgacgcgc	gctgggatct	ggcaggcttg	45240
ttcgatccgg	acccggaacg	ggctgggaag	acctacgtgc	gggaaggggc	cttcctcacc	45300
gaogccgac	ggttcgatgc	gggtttcttc	gggatttccc	cgcgtagagg	gttgccgatg	45360
gatccgcagc	aacggctgtt	gctggagctg	tccctgggag	ccattgaacg	ggcagggatc	45420
gatccgggtt	cgctgagggg	gagtcggacc	ggtgtgttcg	cggggctgat	gtaccacgac	45480
tatggcgccc	ggttcgccag	ccgagccccg	gaaggtttcg	aggggtatct	cggcaatggc	45540
agtgcgtgga	gtgtcgcgct	gggcccggat	gcgtactcgt	ttggtttcga	gggtcctgcg	45600
gtgacgggtg	atactgcgtg	ttcgtcgtcg	ttggtggcgt	tgcattttgg	gggtcagtcg	45660
ttgctgtccg	gcgaatgcga	tctcgccctt	gccgggtggt	tgacgggtgat	gtcgaagccc	45720
gggacgtttg	tgggaattctc	cgtcagcggg	ggcctggcac	cggacggggc	gtgcaagtcg	45780
ttcgcggaga	gcgcggacgg	taccgggttg	ggtgaggggt	ctgggtttgt	gttggtggag	45840
cggttgtcgg	atgctcggcg	gaatgggcat	cgggtgttgg	cgggtggttcg	tgggtcggcg	45900
gtgaatcagg	atggtgcgtc	gaatggcttg	accgcgccga	atggtccctc	gcagcagcgg	45960
gtcatccagc	agggcgttgg	gagtcggggt	ctgtcgggtg	ccgatgtgga	tgccgtggag	46020
gcgcatggga	ccgggaccag	gttgggtgat	ccgattgagg	cgcaggctct	gattgctacg	46080

tatgggagcg atcgtg cggtcggccg ttgtggttgg ggtcga aa gtccaacatc 46140  
 ggtcatacgc aggcggcgcc ggtgttgcc ggtgtgatca agatggtgat ggcgatgcgg 46200  
 cacgggcaac ttccgcgcac gctgcacgtg gatgcaccct cctcgcaggt ggattggtcg 46260  
 gcggggaggg tccagctcct gacggagaac acgccctggc ccgacagtgg tcgcccctgt 46320  
 cgggtggggg tgtcgtcgtt cgggatcagc ggcaccaacg cgcacgtcat cctggaacag 46380  
 tccacggggc agatggatca ggcagcggag cggattcga gtcctgttct ggatgttccg 46440  
 gtggtgcccgt ggggtggtgc gggcaaaaca cccgaagcgc tatccgcca ggcggcaacg 46500  
 ttggcgacct atttgacca aaatgttgat gtctccctc tggacgttg gatttcgctt 46560  
 gcggtgaccc gttcggcgct ggatgagcgg gcggtggtgc tggggtcgga tcgtgacacg 46620  
 ttgttgtctg gcctgaatgc gctggctgcc ggtcatgagg ctgctggcgt ggttacggga 46680  
 cctgtcggga ttggtggccg gaccgggttt gtgttcgcc gtcaaggcgg tcagtgggtg 46740  
 gggatgggcc gccggttgta ctcgagttt ccggcgcttc ccggtgcttt cgacgaagca 46800  
 tgcgccgagc tcgatgcgaa cctggggagg gaagtcgggg ttcgggatgt ggtgttcggc 46860  
 tccgacgagt ccttgcctga tcggactttg tgggcgcagt cgggtttgtt cgcgttgacg 46920  
 gtcggtctct gggaattgtt gggtagctgg ggtgttcggc ccagcgtagt gctggggcat 46980  
 tcggtcgggg agctagccgc ggcgttcgcc gcaggtgtgc tgcgatggc ggaggcggct 47040  
 cggctggtgg cgggtcgtgc gcggttgatg caggcgttgc cttctggcgg tgccatgctg 47100  
 gcggtgtccg cgaccgaggc ccgagtcggc ccgctgctcg atggggtgcg ggatcgtgtt 47160  
 ggtgtcgcag cggttaacgc tccggggctg gtggtgcttt ccggtgaccg ggatgtgctc 47220  
 gatggcattg ccggtcggct ggacgggcaa ggtatccggt cgaggtggtt gcgggtttcg 47280  
 cacgcgtttc attcgcacgc gatggatccg atgctggcgg agttcgccga gctcgacgg 47340  
 agcgtggact accggtctcc acggctgccg attgtctcga cgtgacgg aaacctcgat 47400  
 gacgtgggcg tgatggctac gccggagtat tgggtgcgcc aggtgcgaga gcccgccgc 47460  
 ttcgccgacg gtgtccaggc gcttgtggac caaggcgtcg acacgattgt ggaactcgg 47520  
 ccggacgggg cgttgtcgag cttggttcaa gagtgtgtgg cggagtcgg gcgggcgacg 47580  
 gggattccgt tgggtcggag agaccgtgat gaggtccgaa cgggtgctgga cgctttggcg 47640  
 cagaccacaca ctcgtggtgg cgcggtggac tgggggtcat ttttcgctgg tacgagggca 47700  
 acgcaagtcg accttccac gtatgccttc caacgacagc ggtactggct ggagccatcg 47760  
 gattccggtg atgtgaccgg tggtggcctg accggggcgg agcatccgct gttgggtgcc 47820  
 gtggtgcgg tcgcgggcgg cgatgaggtg ctgctgaccg gcaggctgtc ggtggggacg 47880  
 catccgtggc tggcggaaca ccgcgtgctg ggcgaagtcg tcgtccccgg caccgcgttg 47940  
 ctggagatgg cgtggcgggc cggtagccag gtcggttggt aacgtgtgga ggagctcacc 48000  
 ttggaggcac cgtggtcct gccggagcgg ggcgctgcgg cgggtgcagt ggcggtgggg 48060  
 gctccggatg aggcggccg gcgcagtttg cagctctatt cccgaggcgc tgatgaagac 48120  
 ggcgactggc ggcggattgc ctccgggctg ttggcccagg ccaatgcggt gccgcggcg 48180  
 gattcgacgg catggccgcc ggacggcgcc gggcaggtcg atctggcgga gttctacgag 48240  
 cgctcgccg agcgcggctt gacctacgt ccggtattcc aagggtccg cgcgcgatgg 48300  
 cggcacggcg acgatatctt gccgaattg gccgggtcac cagacgcctc gggtttcggc 48360  
 atccaccgg cgctgctgga cgctgcactg cacgcgatgg cgcttggtgc ttcgcccagc 48420  
 tcggaagcgc gtctgccgtt ttctggcgt ggcgccagc tgtaccgcgc tgaaggagca 48480  
 gcgcttcggg tacggtctct gccgctggg tccggtgcag tctcattgac gttggtggat 48540  
 gccacagggc gacgagtcgc tgcggtggaa tcgctttcga cgcgaccggt ctccaccgac 48600  
 cagatcgggt cgggtcgcgg cgatcaagag cggctgctgc acgtcagtg ggtaaggctc 48660  
 gctgaatctg cggggatgtc tctgacctc tgcgcggtgg tcggtttggg cgaaccggag 48720  
 tggcacgtg cgtgaagac cactggtgtc caagtcagat cccatgcgga ccttgcttcg 48780  
 ttggccaccg aggttgccaa gcgggggttc gctcctggtg cggtcacgt cccgtgcccg 48840  
 cgaccccgag cgatgcagga gctgccgacc gccgcgcgaa gggcgacgca acaggcgatg 48900  
 gcgatgctgc agcaatggct tgccgatgac cggttcgtca gtacggcct gatcctgctg 48960

acgcacatcg	cggtctccgc	agttgctgga	gaagacgtgc	tcgacctggt	acacgcgcgc	49020
ctgtggggct	tggtccgcag	cgcgcaagcg	gagcaccgcg	accgattcgc	cttgatcgat	49080
atggacgacg	agcgagcatc	gcagacggca	ctcgccgaag	cgctgactgc	gggagaagcg	49140
cagctcgcg	tgcggtcggg	agttgtgctg	gcgccccgc	tcggccaggt	gaaggtgagt	49200
ggaggtgaag	cgttcaggtg	ggatgaaggc	accgtgctgg	tcaccggcgg	aaccggcggg	49260
ctcgggggcc	tgctcgcacg	ccatctggtc	agcgcccacg	gtgtgcggca	cctgttgctc	49320
gcaagtcgcc	gtggtctggc	ggcgcccgga	gcggatgagc	tggtggccga	gctggagcag	49380
gccggcgccg	acgtcgcggt	cgctcgctgc	gactcggcag	atcgggactc	gcttgcgcg	49440
ctggtggcgt	cggtgcctgc	ggaaaacccg	ttgcgggtgg	tggtgcacgc	cgccggtgtg	49500
ctggatgacg	gtgtgctgat	gtcgatgtcg	ccggagcgct	tggaacgcgt	gttgcgccc	49560
aaagtggatg	ccgctgggta	cctgcacgag	ctgactcggg	aactcggtct	gtcggcgctc	49620
gtgttggtct	cctcggtcgc	gggcctgttc	ggcggtgcgg	ggcagagcaa	ttacgtgccc	49680
ggcaacgctt	tcctggatgc	cttggcgcac	tgccggcagg	cccaggggct	gcccgcgctg	49740
tcgctggcct	ccgggctgtg	ggcgagtatc	gatggaatgg	cgggcgacct	cgctgcggca	49800
gatgtggagc	ggctgtcgcg	ggcaggcatt	ggcccgtttt	cggcaccggg	agggctggcc	49860
ttgttcgacg	ctgccgttgg	ctcgagcga	ccgttgctgg	caccggtgcg	actggatgtc	49920
gaagcactgc	gtgtgcaggc	ccgatccgtg	cagaccgcga	ttccggaaat	gctgcatggc	49980
atggcaatgg	ggccaagccg					50000

<210> 3

<211> 25360

<212> DNA

<213> Saccharopolyspora spinosa

<400> 3

ccgcactccg	ttcaactcca	gggttgagcc	gttgcaacga	cggttgcccg	gattgtcgga	60
gggcgaacgt	cggcagcaag	tgctccagcg	cgctccgcgc	gatatacgcg	tggtactggg	120
gcacggcagg	tcgagcgatg	tggaacatcg	gaagcctttg	gccgagctgg	gtttcgactc	180
gctgacggcc	atcgaaactc	gcaaccgtct	cgctaccgc	accggactgc	ggcttcccgc	240
gacgctggcc	ttcgaccacg	gcaactgcgg	ggcaactcgc	cagcacgtgt	gcgcgcagct	300
aggcaccgcg	accgcgcgcg	caccgaggcg	aaccgacgac	aacgacgcca	cggagcccgt	360
gaggtcgctc	ttccaacagg	cgtatgcggc	tgcccgata	cttgacggga	tggtatttgt	420
gaaggtcgct	gcccagttgc	gaccggtgtt	cggttcgcct	ggcgagctgg	aatccctgcc	480
gaaaccgctc	cagctttccc	gtggtcccga	agagcttgcc	ttggtgtgca	tgccggcgct	540
gatcgggatg	ccgcccgcac	agcagtacgc	gcggatcgcc	gccgggttcc	gcgatgtgcg	600
ggacgtttcg	gtgatcccga	tgccctggatt	cattgcggga	gaaccgctgc	cgctccgcat	660
cgaggtggcg	gttcggacgc	aggcggaggc	ggtgctgcag	gaattcgccg	ggggtcgttt	720
cgtactggtc	gggcattcct	ccgggggctg	gctggcgcac	gaggtagccg	gtgagctgga	780
gcgtcgcggg	gtcgtcccgg	ccggggctgt	actgctggac	acctacatcc	ccggtgagat	840
cacgccgagg	ttctccgtgg	cgatggccca	ccggacgtat	gagaagctcg	cgactttcac	900
ggacatgcag	gatgtcggtg	tcaccgcgat	ggcggggtac	ttccggatgt	tcaccgagt	960
<hr/>						
gactccgacg	ccgatcgggtg	ctccgacgct	gttcgtgcgg	accgaagatt	gcgtcgcaga	1020
ccctgaaggg	cgcccggtgga	cagatgactc	ctggcgccca	gggtggactc	tcgcggatgc	1080
cacggtccag	gtgccggggc	accacttctc	gatgatggac	gagcacgcgc	ggtccaccgc	1140
acaggcagtc	gcgagttggc	ttgacaaact	caaccagcgc	accgctcggc	aacgctgacg	1200
ggcgctcctt	taggaccttc	tgggcggcac	cgccaccccc	ggcggtgccc	ccttccgtgg	1260
tccaggctcg	ccgatcttga	cggcgcacga	tgccggcgac	gcgcgctgat	cgtgattccg	1320

ctgccgctcg tggcca cctggcgaat catgtccttt cgggcgt caaacgaatt 1380  
 cgtccgagcc cgcattccga ggtgaggggc acccttgggt ggctgagccg ctcaaggggtg 1440  
 cccctcacct cgaaattcgt ccgatttggg cggtagacgc aaccccggtg ggcgtggtgc 1500  
 gtctttcttg ttgacagagc ggtgagaagc cgctgacaca cctgagagga aaaggggagc 1560  
 atgatgctca agcgccaccg ttgacgacc gccatcaccg gccttctggg gggagtactg 1620  
 ctggtcagcg gctgcggaac cgcgcgcga cttcagtcct cgcggcgcc cgggcatgac 1680  
 gcgcgcaatg ttggtatggc ctggggcggg ggcggcgggg acatcggcac gtcgaactgc 1740  
 tcggaggccg atttctcgc caccgcgaca ccggtgaaag gcgaccccg cagtttcatc 1800  
 gtggcgtacg ggaaccggtc ggacaagacc tgcacgatca acggcgggcg gccgaacctc 1860  
 aagggcgtgg acatgagcaa ctgcgcgac gaggacctgc cggtcgagga cgtgcggctt 1920  
 cccgacgcgc ccaaggaatt caccctccag ccgggtcaga gcgcgtacgc cggcattggc 1980  
 atggtcctgg ccgacagcgg cgaaccgaac gcccatgtcc tcaccgggtt ccagtcctcg 2040  
 ctgccggaca tgtccgaggc ccagccggtc aacgttctcg gcgacggcaa cgtgaagtgc 2100  
 gccgcgaagt acctgcgagt cagctcgctg gtgtctaccg cagacgagct gcgctaaaac 2160  
 ccatgtgagt cccgcagatt cgacctcgcc gtgcggcgcc tccggcgaag cgtccgtacg 2220  
 tttgtcgttg tgaccagcgt tgttcacgtc cgggcgcgac gctggtacat actcaggcgt 2280  
 ctccggcgcc tccaacgggg cctggcatcc ggggcgctcg agtgcggcgg cgtgacgcg- 2340  
 ttctctgtcg ggcgttgtca cgcgcgcggc ctgcaaccgg tcccgcgccg tcggagccgg 2400  
 tgggtccagcg cgggtgtggcg gcggccggag ccgacggtgc gcaccgcctg cccgagggcc 2460  
 tttttcgaac cgacgaggac cagaccttc ttggcccggg tgaccgcctg gtagagcagg 2520  
 ttgcgctgca gcatcatcca ggcgcttggt gtcaagggga tcaccacgca cgggtattcg 2580  
 cttccctgcg aacgatggat ggtcaccgcg taggcgtgga ccagttcgtc gagttctgtg 2640  
 aagtcgtagt cgatgtcctc gtctcgtcg gttcgcacgg tcatggtctg tgcttcgttg 2700  
 tcgagggcgg acacgacgcc ctgcgtgccg ttgaacacgc cgttggcgcc cttgtcgtag 2760  
 ttgttgcgga tctgcgtgac cttgtcgccg acgcggaaga tccgtccgcc gaaccgcgc 2820  
 tctggcaggc cttccctggc cggggtgatc gcttctgca acagctggtt cagcgcgcct 2880  
 gcacctgcgg ggcctcgatg catcggggcg aggacctgca cgtcgggtcg cgggttgaac 2940  
 cggaacttcc gcggaatccg gcgggcgacg acgtcgacgg tgagctcggc ggtcggttcg 3000  
 ctttctcta cgtggaacag gaagaagtcg gtcagcccgt gtgtcagcg atagtcgccg 3060  
 gcgttgattc ggtgcgcgtt ggtcaccacc ccgactcgg cggcctgccg gaacacctcg 3120  
 ttgagccgca cgtgtggaat cggggtgcca ggggcgagca gatcgcgag tacctcaccg 3180  
 gctccgaccg acgggagctg gtgcagtcg ccgaccagca gcaggtgcgc gccgggcgcg 3240  
 atgccttg ccagtttgtt ggctaacagc aggtcgagca tggacgcctc gtgaccaacg 3300  
 acgaggtcgg cgtccagcgg gttgtcccg tcgtaggcgg cgtcccgcc cggctggagt 3360  
 tggagcaggg ggtgcacggt cgcgcgctcg tgtccggtga gctcggtcag ccgcttcgcc 3420  
 gctcgtcccg tcggcgcgcc gaggatcacc ttggcctttt tcgcctgagc taatgcgatg 3480  
 atcgaccgca cgggtgaagct cttgccgcag cctggacctc cggtagacac ggcgaccttc 3540  
 tcggtcaggg ccagcttgac ggcgcgctcc tgcgcctcgg cgagttcggc accggtagcg 3600  
 cggcgcaacc agtcgagggc cttgtgcaa tcgacgtcgg cgaagacggg catccggctc 3660  
 gcgctggtgt tcagcagccg ggacagctgg ttggccagg cgacttcggc gcggtggaag 3720  
 ggcacgaggt agatcgcgac cgtcggcacc tcgtcgtcat cggtagggat ctctcgcgg 3780  
 accacacctt cctcggtgac gaggtcggcg aggcattoga tcaccagccc ggtgtcgacg 3840  
 gcgaggatct tcaccgcctc ggcatcagc tcgttctccg gcaggtagca gttgccgtcg 3900  
 ccggtggact ccgacagcgt gaactgaag cccgccttta cccgctgcgg ggagtcgtgc 3960  
 gggattccca ccgctttggc gatggtgtcg gcggtcttga aaccgattcc ccacacgtcg 4020  
 cctgccagcc ggtatggctc ttccttgacg gtccggatcg cgtcgtcgtg gtactgcttg 4080  
 tagatcttca ccgccagcga ggtcgagacg ccgacgcctt gcaggaagat catcacctcc 4140  
 ttgatgcct tctgtcctc ccacgcgtcg gcgatcagct tcgtcgcgtt cgggcgcgagc 4200



ttggggacct	cgatcagccg	cgcggggttc	tgctcgatga	cgtcgagcgc	ggcgacgccc	4260
aagtggtcga	cgatcttctc	ggcgagtttg	gggccgatgc	ccttgatcag	gccagacccc	4320
aggtagcggc	ggataccttg	cacggtcgca	ggcagcacgg	tcgtgtagtc	gtcgacgtgg	4380
aactgccgcc	cgtactgggg	gtgcgacccc	caccggccgc	gcatgcgcaa	cgcctcgccg	4440
ggctgcgcgc	ccagcagcgc	gccgacgacc	gtcaccaggt	caccgccccg	gccggtgtcg	4500
atccgcgcga	cggtgtagcc	gctctcctcg	ttggcgaacg	tgatccgctc	cagcgtgccc	4560
tccagcaccg	cagtccacgt	ggccgactcc	cgtccttttt	ccaccgacaa	cacgtatcac	4620
gaacggctgt	caagcaaacc	ggcggtcacc	acatgcagcg	gcatctcccc	aacgcctcgg	4680
gctccggcgt	cagcgggtgg	gcgttcgcga	tgcttgggtg	cggccgggtg	gagttgtaga	4740
ttttttcgtc	ctcgcgcgag	gcctggagta	ggtgccgctg	gctccagatc	aggggaaagc	4800
gccgggataa	accggcttga	cggaggagat	ggaagagctc	tacgtcgaat	ggccagcgac	4860
ccacgacgac	cccaggtcat	gcgtcgacga	cctgcgaggc	cgtggcgaag	cgttgacagg	4920
gggcacgtgg	gggcgtgcga	tcaggtcgcg	gcgcgggctc	agcgccgtgt	ccggacgcac	4980
cagcagccgc	gacggactgg	ttccgctcct	cgagcgcggt	gtcctcgatc	gtcttttttcg	5040
agctcgatgc	cggccgcgtt	cgcgagggtg	ccagtcacgt	cggtcggcag	tcgtggtcgc	5100
agtctgcctt	ttgaacatct	ggtcaattca	gctttgaatg	atcgttcggc	gagtattgtc	5160
ctgcccgcga	attgtttgct	cgccgtgcgc	cctgggacag	gagcgcgggc	ttcggggacg	5220
gcgtgctggc	gaaggtgtcg	attggctcgg	acggccggac	caggatcgcc	gacgtcgtga	5280
tcgccacgag	cgtcggcgat	gcccggtact	atctggagca	ccagcgagtg	gacaacgact	5340
tccacgggcg	cggtgcgcctc	ctgatcatga	atcaagaact	gagaccgtca	gggtgaacgt	5400
gccgtcgtgg	ggcactgtga	ggatgtgcat	cacacggccg	gggtaggtga	tcgaggaggc	5460
cctgatcgcg	ggctgccgcg	gctgtgccca	gtcgaatcgc	agcaccacct	ggccgcgggt	5520
gagcgtcagc	ccctgcgcgt	cgagtgcgac	tttgtcggag	tccttcacca	ccaacgggat	5580
ctgctcgggt	gccggtgcgt	tcgcctgcac	cgacctgggt	atccggtcgg	gccgcacggt	5640
cacatcggtg	acgaccgagc	cattcgcggt	gcggtagcgg	aaccggtatg	gctcgccgcc	5700
caggaactcc	gcctgctgcg	gtccgttcgc	gtccggcacc	tggcccggga	agaccgtcgc	5760
ccagcaggcg	tggtcgttgt	tgttcatoga	ctggacgacc	gtgcccgoga	ccgggtgcca	5820
caggaacgtc	agaccgggtg	ggcccagcga	ggtcgcgcgc	ctcgtgccga	agtagccgcc	5880
gaggtagaac	tgcggcctgc	gtaggaacag	gaagtcttgc	ccctggtcgc	tgcggtatctc	5940
ggtgaagtgc	gcctgttggt	atggaagttt	cgcgatcgcg	gccgtgcgct	gggcgcgggt	6000
cggatagcgc	tcgccgtaga	tcgcgtgggt	gaggatgcgc	ggtgaggtgt	cctgtttgac	6060
gagtgcgggc	accggtgccg	ggtcctcggc	ccaggccgct	cgtgcggcgg	cgaggtcttc	6120
gcggctggac	aggaacgcgc	ccaagttcgg	caccactggc	acgaactggg	tgttcagcgc	6180
ggtacgttcc	gggtccggtc	gcacgtcggc	gtagaaccgc	gccgaggtgc	gggagttggg	6240
ggcgacgttg	gcgaaccagc	cggagccgct	gggttcgcgc	agcaggttgt	agctgagcca	6300
gtcgggtgtac	ttccgcgcca	tcgtgacgag	gttggcggtg	ttggactggc	gccacgcgtc	6360
ggccatctcc	ggcagcatca	cctcgaagtt	gtagttgggt	tcgcgcgcgg	tcaactcgta	6420
gaagtagccc	gaaggactct	gcccgttcgc	ggcaaggcgc	gcgaacgcct	cgggtcaactg	6480
ctgccgcagc	gcggggtcgg	ggtcctgggt	cagcgcgagc	gaggagccgg	cgagcccggc	6540
ggtaacctgg	ttcgcgtagt	ggattgtggg	ctgccaggcg	ctaccgttgc	ccgggttgag	6600
gagccacgtc	atcgctttgc	gcagcgccga	ggtgatctgc	gactggcggt	gcggcaggat	6660
gttcgcgcgc	cgcagcaggg	cgttgggtctt	gctcaggtac	cccaaccoga	agccgggtggc	6720
ggcgagcccc	tgctcggtcg	gcgaataactc	cggccacgag	ccgtcatcgt	gctgcaaccc	6780
taggtagtgt	ccgagccccg	cgtcgagcgc	ggcgagcaga	gtcgcgtcgc	cccggtacgg	6840
gttccacgtg	cgggatttgc	tggcgaacca	ggcgagcgtg	tagacgtgct	cctggacgcg	6900
ggcgttgtag	gacaccgcgc	ggctgcgcca	ccagcctccg	gcgaagaagc	cgtcagagtc	6960
catgtcggcg	accatcggcg	agaccgcggg	gaggtacgac	gcgaatcgct	gctcctcggg	7020
tgccaacagg	cgccggttcg	gtgtgggtggg	cggcggggcg	ggcagtgcca	gtgcgcgggt	7080

cggcagtgcg gcgagcgc cgagagcggc ggccgcggtc atcagcgc gacggctgaa 7140  
 cgtagtcacg ggcctacctc cttgtggccg atcaaccctc acccgctgcg tagccgcacg 7200  
 tcaagatgat aattcgaatt attatgggct tgacgacgcg taggccgacg acgcagaatt 7260  
 cctgccaatt cgtattggca agcgggggtg ctcgtggccc gacggctcac ccagcaggac 7320  
 atcgcccggg tggcaggagt cagccaggcc acggtgtcgc tgggtgctca caaccggaag 7380  
 gacggcaacg tccggatcgc ggccgagacc cgtgcgcagg tactggaggt gattcggaag 7440  
 accggctacg tcgcgaaccc gatcgccgcg aggatgcgcg atcggcacaa ccgcacctc 7500  
 ggctgtttca cctacgaggc ggtgttcccg agcaccacg cgaacttcta ccagtcgttc 7560  
 ctcgaaggca tcgaggaaca ggccgaggag gtcggctgtg acctgttgtt gttcaccagc 7620  
 gccaaaggcca ccggggagcg gcggcggtt ttcggcgacg acagccgggt gcggctcgcc 7680  
 gacggcagtc tgctgctcgg tcgcacggtc gaccgcgacg acctgaccca gctgctcgcc 7740  
 gaaagcatcc cgtacgtctc catcgacga cgcgacgacg cgggcgggtcc ggtgccgcac 7800  
 gtcggggccg actaccgcac cgcggtgcga gacctggtg accgcgcggt cgcgctcgcc 7860  
 caccgcgggt tcgcgtacgt ggggtctggt gggggcgcgg agtcgtccgc ggatcggctg 7920  
 cgaggcttcc gcgaagccgt tgccgcacat ggctccaag ggatgcatgt ggagacccca 7980  
 cagctcgatc agctgcgcga agcgggcgtc accgctgtgc tcaccgaaga ggtgtcggac 8040  
 ggggcgcgcg tcgtgctcgc cgggcgcgaa cgcgggctct ccgtgccggg cgacctcgcc-8100  
 gtgctctcgc tcggtgcgcg taccggctcg gcaccggacg acgacgtggc gctcaccggt 8160  
 ttccgcatcc ccaggcgcga gatggggcgc cgggcgggtg aggcgctgac cgaggtgctc 8220  
 gaaaacggca ccacaccgca agaactgctc ccgtgcgagt tcgtcgaggg ctgcacgctg 8280  
 ggcgaccac gcctttgacc aggaggaact gttgctcgac cacaccacgg acgttgctgt 8340  
 cgttgccggc ggactcggcg gtgtcgcgcg cgcactcgcg ttgctgcgcg cgggcgcgcg 8400  
 ggtcgtgctc accgaggagt acgactggt cggcgccag ctgaccagcc aggcgctgcc 8460  
 gccgacgag cacagctggg tggagcgctt cggcgtcacc gcgagctacc gggcgttacg 8520  
 cgacggcatc cgcgactact accgcgcga ctaccgctg acccgcgcg caccggcgtg 8580  
 gcgggagctc aaccccggtg cgggcaacgt gagcaggctc tgccacgagc cccgcgtcgc 8640  
 cgctcgcggtg atcgacgaga tgctggcgcc gttccgcggc agtggcaggc tgaccgtgct 8700  
 gcagccgtac cggccggtgg ccgcgcacaa cgacggcgac cggatcgtgt cggtgaccgt 8760  
 tgccgaccgc gacaccggtg aacagatcga gctctccgcg ccgtacatcc tggacgcgac 8820  
 ggagacgggt gaactgcttc cgttgtccag cacggagtac gtcaccggct tcgagtccac 8880  
 tctggacacc ggcgagccga gtgcgcgcga cgtcgcgcag ccggcgaaca tgcaggcgg 8940  
 gtcggtgtgc ttcgtggctg accacgtcga cggcgaccac accatcgaca aaccggcgcg 9000  
 gtacgacttc tggcgcgcgt accagccgga cttctggggc gaccggatgc tgtcgttccg 9060  
 ctccccaaac ccgcgcacgc tcgcgatctc cgaacgtacg ttcacccga acccgacga 9120  
 cgaccgcgtc ggcgtcgtgt cggaccagcg gctcagtgc ggtgacagca atctgtggac 9180  
 gttccggcgc atcgccgcgc gtcgcaactt cgtcgagggg gcctacgaca gcgacatctg 9240  
 cctggtgaac tggccgatca tcgactactt cgagtgcgcg gtgatcgacg tgccggacgc 9300  
 cgacgcgcac atcgccgcgg cgcgggaact ctcacgttcg gtgctctact ggctacagac 9360  
 cgaggcgccg gcgccagacg gcggcaccgg cttccccggc ctccgcctgc gcggcgacgt 9420  
 caccggcagc gcggacggtc tcgcgcaggc gccgtacatc cgcgagtcca ggccgcatcag 9480  
 ggccgagcac acgatcgtcg aacaggacct ctcgctcgcc gtgcgcggcg acaagggtgc 9540  
 ggtgcagcac gcgcagccg tgggtgtcgg catgtaocgt atcgacctgc acccctccac 9600  
 cggtggcgac aactacatcg acgtcgcgag ctgcccgttc gagatccgcg tcggcgcgct 9660  
 gatccgcaa cgggtggaga acctgctacc cgcgggcaag aacatcggca ccaccacat 9720  
 caccaacggt tcccaccggc tgcaccagt cgagtggaac gtcggcgagg tcgcgggcgc 9780  
 gctcgtgccc ttctgctgg cgcaccgagt caccctcgc gcggtgcgca ataccctgg 9840  
 cctgctcgcg gacttccagc agtgtctgga acgcgacggg gtcgagctcc gctggccgga 9900  
 cgtgtccggc tactgacga gggagacgaa aatgacaaag ctgtcacgac gactcacggc 9960

actcatcgtc	gcagggctgt	tcgccctcac	cggctgcggt	ggtggatcaa	ccgcacagtc	10020
cggaccgaag	tcgctgcgca	tgaccgtgtg	gactgccaac	gcggcgcatc	tcaagctgct	10080
caacgacatc	gccgccgagt	acaaggcctc	gcacccggac	atcgccgaga	tcaagttcga	10140
ctcgggtgcc	gccgacggct	acaccaccac	gctcaccacc	cagatcgccg	gcggtaacgc	10200
gccggacctg	gcctggatcc	tggaggagtc	ggcaccggac	ttcgtggcgt	ccggtgcgct	10260
cgccccggtg	cgcggaaga	tcgagaaggc	cgacgagctc	gtgccgtccg	cgacgaagct	10320
gtgggagaag	gacggcgaac	tgtacgccta	cccgttctcc	acctcgccgt	tcggcttggt	10380
cgtcaacacc	gacctggtga	agggcgccctc	ggcggactgg	acctgggacc	aggcgatcgc	10440
ggctgcctct	gcgtcgggcg	ccgcctccgg	caagggcggc	ctggtactgc	cggacttcaa	10500
gtaccagaac	tgggcagtgc	tgtcctctat	ctggcgcggc	tggggagctg	atgcgtggag	10560
cgcggaagg	cgctcggtgc	ggttctccag	cagcgagatg	aacgacgcga	tgtccttctt	10620
gcacaaggcc	atcttcaccg	acaaggcgat	tccgggcccc	ggcacgacgg	tggacttctt	10680
cgccggcgac	gcggcgatgg	cgatcggcca	gatctcccgg	tccagtgcgt	tgaaggacgc	10740
gaagtccggc	tggacgtgc	tgccgctgcc	ggccggtccg	aagggtgact	acgcggtgat	10800
cgggcaggcc	gggatcggtg	tgctgaagca	gtcgcacaa	gtcgacgccg	cgacggactt	10860
cctcgccctc	ctcaccaacc	agaccaactc	cgccaagctc	gcgcagttct	cccccccggc	10920
gcggtcgctg	ctgctcaacg	cggagacgct	cgccaagagc	aaccggttga	tcaaggccga	10980
gcaactgcag	tcgctcgctg	tcgacggcat	caacaagggc	gtcgtgaagc	cgagccacaa	11040
gggtcaggag	gagctgaacc	agacgatccg	cgccgcgctc	gaccggttgt	ggaagccgga	11100
cggaacgtg	cagaacgtgc	tgaacgacgt	gtgcaccaag	atcaaaccgc	tgtcggagaa	11160
caagtgcgg	cggtcgcaca	ctccacgcgg	gccgtagccc	gcaaggggac	gtcgtactgg	11220
acgcagcgca	ggcgcgacaa	tctggttgg	tacctgttcg	tcgcgcccg	gctgctcgcc	11280
agcatcgctg	tcgtgctggt	gccgctggcc	ttggtcggct	ggtacagcct	caacgagtgg	11340
aacgtgctcg	ccggcacgtt	cgagttcgct	ggcgcgcaga	actaccaaca	gctgctcgcc	11400
gacgagaagc	tgcgcgactc	gctggtggcg	accacttggt	tcgccgccgg	cctggtcgctg	11460
ctcaacctgt	cactagcgct	gctgctggcc	gtgctgctca	accagaagct	gagcggcacc	11520
acggtcttcc	gcacgttggt	cttctctccc	gtcctcgctg	cgctggtggc	gtggacgctg	11580
gtgtggcagc	tgatactgca	gccggagggc	agcgtcaacg	gactgctcgg	gttcttcggt	11640
gccgacgggc	cgaactggct	gcgcggtgag	tcgaccgcga	tggctcgggt	gatcgtcgctg	11700
caggtgctca	agaacgtcgg	cctgaacatg	gtgctgttcc	tcgcggcgct	gcaaggcgctg	11760
ccgcagccgt	tgtacgaggc	ggcgaagatc	gacggggcgg	gtgcctggac	ccggttccgc	11820
cgcatacct	tgccgttgat	cagcccagcg	atcctgctca	cgtcgatcat	caccatcgctc	11880
ggctcgctgc	aggtgttcgc	gcagatcgcg	gtgctcacgc	agggcggtcc	gggcacgtcc	11940
acgaccgtac	tgatctacta	cctgtaccag	caggcggtcc	agttccacca	cttcggctac	12000
ggcgcgacga	tctcggtact	gctgttcgct	atcgtcgccg	cactcaccct	gctgcagtgg	12060
cagatgcgcc	ggaagtgggt	gctgcatgag	gcttaggggc	aagatcacgc	tatacgggct	12120
gctgtgcctg	ctgtgcgtgc	cttctggtgt	cccaacctgg	tggatgatca	cctcgtcgat	12180
gaagcccatc	agcgagatct	tctccacctc	gccgctgccg	tcggagtgga	cgttctccac	12240
ctaccggcag	gtgttcgaga	tgcagccgtt	cgcgcgacag	tactggaaca	gcctctacat	12300
cgcggtgatc	gtcacgacag	gcacaatggc	ggtggcgcg	atggccggat	acgcgttcgc	12360
gcgtatccgg	ttcccgggcc	agaacgtgct	gttcgtgggc	gtcctgatcg	gctgctcat	12420
cccgagcgag	gtcaccatcg	tgcgctgtt	caagatgttc	cagtcgctcg	gcctgaccaa	12480
<hr/>						
cacgcactgg	ccgctaataca	tcgtgccgat	cttcggggcg	ccctgtgtgc	tggcgatctt	12540
catcatgcgc	cagttcttca	tcgccctgcc	gtcggagctg	gaggaggccg	cgcggtgga	12600
cggcctcgga	cgagcgggca	tcttctaccg	ggtggccctg	ccgctgtccc	ggccgcgct	12660
gggcgcgctc	gcgatcttca	cgttcctgca	ctcgtggaac	ctctacctgg	agccgatcgt	12720
ctatctgtcc	acaccggaca	tgtacacgtt	gccgcaagcg	ctcacgcagt	tcgtcgacaa	12780
ctacggcggc	ccgatgtgga	acgtgcagct	cgccgcgcg	accatgaccg	cagtgccgat	12840

tctcgtcgtg ttcgtg cgacagcggca gttcatcgag ggact acaccggggg 12900  
taaggggtga gtcccgcgat cgacccgag atcgccgcgc tggtagacct cgccagacca 12960  
ccggtgacgc cagcgggaat cgacgcggtc cgggcgggag gacgcgtcgt caccgatgcc 13020  
gagctgacgc gcggtggcac cgtgacgttc gccgatgccg atgccgatgg cgtgcccggtg 13080  
cttgtgctgc gccacgcggg tgtgccacgt cttccgggtgc tgcacctgca cggcggcggg 13140  
atggtcgcgg gcaccggcg caccgacctg cacgtgctgg ccgagtgggt gtcggagctg 13200  
ggtgtcgtgc tgggtgcgcc ggagtaccgg ctcgcacccg agcaccgcga cccggcgccg 13260  
tcgcaggact gcttcgggt gctggagtgg atgtcccgca acggcttcgg cccgcctgtc 13320  
gttgccgggtg catcgcggg cggcgggctg gcggctgcgg tgacgtgat ggcccgcgac 13380  
ctcgcggtgc cgccgatctt gccgtcggcg cgcaactgct cgaacgcgtc cagcgtctcc 13440  
gccaacggcg tgctaccgcg ccagtgcagc aggtagacct ccaggtgatc ggtgcccac 13500  
cgccgcagac tgcgttcgca cgccctgcag gcgccacgcc gaccgcggtt gtgcggatac 13560  
accttgctga ccaggaatac ctcgtcgcgg cgcccgcgga tcgcggcacc gacgacctcc 13620  
tcggcgcccc cgctgccgta catctcggcg gtgtgatca gcccaggcc gaggtccagg 13680  
ccgcgcgcga gcgcggcgac ctccgcgcgg cgctgccac gccgtcccc catgccccag 13740  
gtgccctgcc ccagcaccgg caattcctgc ccgccacgca gttgaaggcc tcgcatgcac 13800  
actccgtttc agtcagcgt gctgcggcag caggccacc agctcgaagc caccgtctc 13860  
ggacggcccc gcggccaacc gccaccggc catctgcacc cgctccgaca ggccggccag 13920  
cccggaccgg ccgctgccga cctcgatggc cgggtcccgc cggtcggggg cgccgtttgt 13980  
gacggcgacg tgcagcgagc cctcgcgga caccagccgc acctgatgt cagcaccggg 14040  
tgcggtgctt gaagcgttcg tcaacgcctc ctggacgacc cggtaagccg cgcgtcaac 14100  
cgacggcgac accccgtgcg ggtcgcggat ctccgaggaa agctccacat ccacaccggc 14160  
ggcgcgcgcc cgacgacca gtcgcgggat cgtgctcaac ccggccccgt gcctgggggc 14220  
gccgagcccc agcgtcgcgc gcatctcgtc cagcgcctgc ttcgccaacc cgcgagccg 14280  
cccggcggtc tccctcgac cgggatcctc ggtggtggcc gccacgccc cgactcgac 14340  
cgcgattaag gtcacgtgat gtcccaccgc gtcgtggatc tcccgcgga tccggggccg 14400  
ctcctcgcg cgcgccgtct cggccttcgc ctgcaactcc gctcgggtgg ccgctccag 14460  
cttcgcgaga ctgtccgca gctgctccc caacgccacc agcgccccca atgcccgtcg 14520  
cgcgccgtgc gccagcagca cgaacgccac ggtcaggatg atcgaccgga ccccagcga 14580  
ctgcatgatc atcaccggtg ccacggcgac cgccgcggtc agcaacaccc agccgaccaa 14640  
catctggatc cgcggtgcg cgcgccccaa cgcgtacaac gccacgacc tggcgccca 14700  
acccaaacca ccggccaacg ctggcatgca cagcagcacc gcgatccgc gccaccgccg 14760  
caacggcagc agcgccgaag ccaccaaac ggacaccgtc gcatagccga acgggtcggg 14820  
cgcgccagc acagctgcc cggaatccc gacgaccagc gcctcaacgc cgagccgaga 14880  
ccagcgacc gaactaccg gaagccgcc gccacgcgt gcgcgatcaa cgcggcctgc 14940  
accggttct gcgcgccgag cttgctcaac accgtcgaca cgtagctctt gaccgtcgcc 15000  
tcggtcagcc cgagccggac ccgatatcg gcgttgacc gaccatccgc caacaactgg 15060  
agcactgcc gtcccgcc ggacaacccc tgcaccagct gcgtctcgtg cgcgttgctg 15120  
gacagactgc ggaacctgg caacaaccgc gcagtgatcc gcggtaccag caccgcccc 15180  
ccggcgcca gatcgtgcac cgccgaacc agcaccgcgg gctccgcgtc cttgagcagg 15240  
aaccggtgc cccgaacc caaggcatcg gccacgtaac cgtccaggtc gaacgtcgtc 15300  
agcaccgccg gcaccggcg attcgccaac gtcgccaagt ccgcgaacgc cgtcagcccc 15360  
tccctgcgg gcatctgcac gtgatgagg gcgacgtcga cgcggtacc gcgcaccgcg 15420  
gccagcaact catccccgtt gccgcctcg gcgaccacc ggatcgagcc gtcgccctcc 15480  
aacaacacc gcaggccac cgcgacatc gctcgtcat cggcaagcac aacgcggatc 15540  
ggccccgc cgctgcagc ccgacctct cggtcacct cctgttccga gggtcgccag 15600  
cagactagc gggccgcga agcgcgagc acacgccacc cgagcgact gtgtcagcg 15660  
cgccgcct tggccttcg ctccgcacc cgcaccgaga tccgaccgg gctaccacg 15720

aacccgaacg	tctcccggaa	cctgcgctcc	aggaaccgcc	ggtacccggc	ctccaagaac	15780
ccggtggtga	acagcacgat	cgtcggcggc	cgcgactgog	cctgggtggc	gaacaggatc	15840
ttcggtgct	tgccgccccg	caccggcggc	gggttggecg	cgaccagctc	ggaaagccaa	15900
ccgttcaacc	ggccggtgga	gatccgggtg	tcccacgaat	ccagcgccgt	gcgagcgctc	15960
ggggccagtt	tcgccaccgc	gcgaccggtc	ttcgccgaca	cgttgaccgc	ctcgcccac	16020
cgcacccgca	ccagctcgcg	gtcgatctcc	ttctccaaact	ggtggcggcg	gtcctcgctg	16080
accagggtccc	acttgttgta	ggcgatcacc	aatgcccgac	cggcctcgac	gaccatagtg	16140
atcacccgca	ggtcctgctc	gctcaacggc	tcggagccgt	cgatgagcac	gatcaccacc	16200
tcggccgcct	cgatcgcggc	cttgggtgcg	agcgacgcgt	agtactcggt	gccgcttgcg	16260
gtcttcacgc	gcttgcgag	ccggcggtg	tcgacgaacc	gccacacctg	gccgtccagc	16320
tccaccagcg	agtccaccgg	gtccaccgtg	gtgcccgga	cgtcgtgcac	caccgaccgc	16380
tcttccccgg	tcagcttggt	cagcaggctc	gacttgccca	cgttcggctt	gccgaccaat	16440
gccacccgcc	gcggcccacc	ggtcgcgccg	aagatctcgc	gcggcgctc	cgggaacacc	16500
tccaggacgg	cgtccagcag	gtcaccggaa	ccgcgcccgt	gcaacgcgct	gaccggcatc	16560
ggctcgccca	accccagcga	ccacagcgaa	tgcacgtcgg	agacgcctcg	ctggtcgctg	16620
accttggttg	cggccagcag	caccggacgc	ttggaccgcc	gcagcaccct	ggccacggct	16680
tcctcggtct	cggtggcacc	gaccggggcg	tccaccacga	gcagcaccgc	atcgggcggtg	16740
tgcattggcca	gttcgcgctg	cgcgggccacc	gacgcctgca	gtcccttcgc	gtccggctcc	16800
cagccaccgg	tgctgaccac	cgtgaagcgg	cgcccgttcc	acaacgcgct	gtagggccacc	16860
cgatcccggg	tgacaccagg	ggtgtcctgc	acgacggctt	cccgacggcc	cagcagccgg	16920
ttcaccagcg	tcgacttgcc	cacgttggga	cggccgacca	cggccagcac	cggctgcggc	16980
ttcgtcggct	cctcgccctc	gaccgcttcg	acctcgctga	actcagccca	ctcagcctcg	17040
tcggaccacg	tgccgtccag	gccttccacc	gactcgctcg	tcacgctcgt	ttccctctcg	17100
aatagtcttg	gctctggtcc	gcgaccagcc	gatccagctc	ggcgatcagc	tcgaccagtt	17160
cgggtgcgac	cgctcggtc	gccacgacca	gcccggcgcg	gcctttgccc	gccggcagca	17220
cgatcggtc	cccgaacaac	acgtcgatcc	tcggtagcag	cctaaggccc	gggccatccg	17280
ggcggcgcgt	gccccggcag	gccaccggca	ggatcagcgc	ctccgacgct	cgcgccaacc	17340
acgcgcgacc	gtgctgcgca	ttggtcacat	cgccatcacc	ccgggtgccc	tccgggaaca	17400
ccccgacgag	accaccggcc	cgcagcacc	gcaccggcgc	cagcagccgg	gcgcggctcc	17460
gcgcaccgcg	gcgcaccccg	atctgcccga	tgcggcgcag	gaaccagccc	agcgcacccg	17520
ggaacatctc	ctgcttgatc	aggaacaccg	cgtggcgcg	gatcattccg	aacagcagcg	17580
gcccgtccat	catcgaactg	tggttgga	ccagtaccac	cgggcccgtg	cgcgggatcc	17640
gctccgcgcc	gtggatccgc	accgggtacg	gcagctgcac	gatccggcgc	gaaatccact	17700
ggccgcaccg	gtgcattcca	ccggacgcgc	cctcgggcag	ggtctcgctg	ctcatggcgt	17760
cgtccgctcc	gccgcgacca	gcaggccacg	gcactccacg	tgcttgcgca	actgcgtgag	17820
cacctccacc	acaccaggt	ccgtggtgtc	gagctccacc	gcgtcgctcg	ccttgcgag	17880
cggcgccacc	ttccgccccg	agtcacgcgc	atcgcggcgc	tgacagtcg	cgtgggtccg	17940
gtccaggtea	ccggcccgtc	cctcgagac	gtcctgcgcc	gtcgggcgct	gcgcgcgggc	18000
gtgcgcggac	gccgtcaggt	agaccttcaa	acccgcaccc	ggggccacca	cgggtccgat	18060
gtcgcgctcc	tcgacgacca	tgccgcctgg	cgacaccagc	gcctcacgga	tcagcttccg	18120
ctggtgggcg	accaactgct	ccgcacctc	ggcgaccgcc	gacaccgcg	acacgcgcgc	18180
ggtcacctcc	ggcccgcgga	tctcccgccc	cacgtccgag	ccgtccagga	agatctcagg	18240
gcccctccgga	tcggtgccc	ccgtcagccg	cgtgcgcgc	accacgcctg	ccactgcggc	18300
cggatccgac	gggtcgacct	cggcgcgag	caccgccagc	gtcacggcgc	ggtacatcgc	18360
tcgggtgtcg	aggtaggtag	cccccaatgc	gaacgccaac	ttgcgggaca	ccgtcgactt	18420
gccggtgccc	gaagggccgt	cgagcgccac	cacaccaagg	agcctggcgt	gtgccacgtg	18480
catgtccttc	ccgagccgga	accgatcggc	ccgacaccgc	gccggaccaa	ccgtccattc	18540
tgccctggtcc	gctcctgccc	atcccaatcg	catccagatc	caatccggac	tgccgggacaa	18600

atacgaaagt tctacgaa aggtcttggg tgtggatgta actgtgccc ccttgccagg 18660  
 gctcggcacg cagcaggact tcaccacccg atccggccac cgcacgagg tgatcaccta 18720  
 ccgggacggc cgcttcgagc tgatcgtctc cgaccacgaa gacccagaca aggtggccgc 18780  
 ctcggtcgcc ttgaccacct ccgagaccag cacgctggcc aacctgctgg gcgccccgca 18840  
 gctggtcgcc cagctcaacg agcagcacag cgaggtcgcc gggatcacca cctggcagct 18900  
 gtccgtcgcc ccgggctccc cctacgccgg ccgaaccctc ggcgacacgg agatgcgcac 18960  
 ccgcacctcc gtgtcgatcg tcgcggtcgt ccgggacggc accgtgcacc cctcaccgcg 19020  
 gccggacttc cagttctccg ccggcgacct ggtggtggtc gtgggcaccg cgaaggggct 19080  
 acgcgcagcc agcgaaatcc tggaaaaggg ctgatccggc ccgtcccga cactttccc 19140  
 gccacctgac cgggacacgc gtcaccaacg gcggtccagc gcgtgccgat cccctaacaa 19200  
 ctccatacgc agggcacacg cgagaggaaa cacgttgcaa gacacggcga tctccctaata 19260  
 agaactgggt gcggttttct tcggtctggg catcctcggg agactcgcat ggaaaatcgg 19320  
 ggtgtcaccg atcccgtct acctgatcgg cgggctggcc ttcggcaccg gtggcctggt 19380  
 cccctgcac ggcacgaac cgttcacca cctcgctcc gagatcggcg tggctctgct 19440  
 gttgttgctg ttaggcctgg aatactcgc ggggtgaactc gtcaccggct tacgccgctc 19500  
 ctggctggcc ggtgtcatag acatcgtgct gaacgcggta ccggtgcgc tgggtggcct 19560  
 gctgctgggt tggggtccgc tcggcgcggt cacaatggcc ggcgtcacct acatctcttc 19620  
 ctccggcatc atcggaagg tgctcgga cctgggtcgg ctcggcaacc gggaaacgcc 19680  
 ggtgatcctg tcgatcctgg tcttcgaaga cctggcgatg gccgtctacc tgccgatcct 19740  
 gaccgcgctg ctggccgggg tcagcttcc cggcgactg acccggtcg ggtgtcgt 19800  
 ggtcgtgate accctggtgc tgggtgctgc gctgcgttc ggcaagtaac tgtccgcgt 19860  
 ggtcgacagc ccgacccgg aggtcttcc gctgcgctg ctggggggcg cgtgctcgt 19920  
 cgcggggatc gctcggagc tgcaggtctc ggcgcggtc ggtcgttcc tgcgggcat 19980  
 cgcatctcc gggtcgaccg cggcgaacgc gaccgcgat ctggagccgc tgcgcacct 20040  
 gttcgcgcgc ctgttcttcg tcgtgttcgg cctcaacacc gatccgagcc agatcccgc 20100  
 ggtgctgcgc gtggcgtgc tgcgtggcgt ggcgacctgc gcgacgaaag tgctcaccg 20160  
 ctggttcgcc gcccgatgc aggtgtcgg ccggatgggt cggctgcggg ccggtgcgc 20220  
 gctggtggcg cgcggcgaat tctccatcgt gatcgccggg ctggccgtcg cctccggcg 20280  
 ggtgccgggc gagctgcgc ccctcgccac cgcctacgtg ctgctgatgg cgatcgtcg 20340  
 ccggtcgcgc gcccgctcg tcgagccggt ggcccgatg ttcctccca gccaggcgaa 20400  
 gaaggcctga ccggaaaagc ctccggtgt cccgagcgga cgggccggag gctttctcaa 20460  
 cgctgcggat aacgcgggc ggtctggcaa acgcaacacc ggtgcgcgaa cagcctcaga 20520  
 ggccactgc gcggtagagc gagccgatct cctggcggt caacggccgg atcgcgccc 20580  
 ggcgctcgtt ggtcagccgc acatcgccga tcgcgtccg caccaggcg tgcaccgggt 20640  
 agccgacgtg cttgagcagc cgccgcagca tgcggttcg gccctcgtgc aacacgatct 20700  
 ccaccagcga ccggttctgg tcatgtcga ccagcttgaa ccggtccacc ttcaccagac 20760  
 cgtcctccag ctccacgcc tcgcgcagct gctgcccgag gtccttcggg atcgaccca 20820  
 gcacctccgc caggtacgtc ttgcgcacct tgtacgagc gtgcacagc cggttcgcca 20880  
 gctcgcgcgc gttggtgatc agcagcaggc cctcgggtgc ctggtccagc cgcgccacgt 20940  
 ggaacagctt gccctcccgt tccgcaggt aatcaccgat gcacggacgc cccttgctgt 21000  
 ccgacatggt gcacaggatg ccgcgcggct tgttcagcag caggtgcgtc agatcgtcgt 21060  
~~tgaccatac ccggttgcgc tccagtgga ccaccgcgt cgtcgggtcg acccgcgcc 21120~~  
 ccagctcggg gacgacctcg ccgtcgacct cgatccgcc ctcgacgatc atctcctcg 21180  
 ccgcgcgcgc cgaggccacc ccggcctggg acaacacttt ctgcagccgc acgcttccg 21240  
 aatcggcagt gctgctgtac ggggaacggg gtcagacgt catcaatcga atccacttcg 21300  
 ggcaacaagg gggcgagcgc cggaactcc ttcaacgacg acagccccag ccgctccagg 21360  
 aacagctcgg tcgtgcagta caggatgcgc ccgctctccg ggtcgggttc ggctcttcg 21420  
 ataaggccgc gcccaccag ggtacggatg acgcgctcga cgttcacacc ccgtactgcc 21480

gcgacccgcg aacgcgtcac cggctgcccg tacgcgatca ccgcgagggg etccagcgcg 21540  
gcacgcgtca gcttcgaacg ctgcccgtcc aacagggtacc gctccacgta cggggcgtag 21600  
ttctcccgcg tgtagaaccg ccagccgtcg ccgacccgcc gcaggtcaac gccccgttcc 21660  
gcctcggcgt agccgtccga gagccgcgcg agcgcggacc ggatccgggc caccggctgc 21720  
tccaggggtg cggccagcag ctctcgcgcg gccggcacgt cgaccacgag cagcaacgcc 21780  
tccagcgcg catccagcgc cgaatcggag gtcagggtcc gcgcactcgg cgcctccacc 21840  
gcctcgggaat cccagcggc cgccgacggc tccaacggtt cctgcggctc agcgggcggc 21900  
ggggtcggcg actcctccgc gacctcgcgc gggcctccg gctccagctc cgtcaccctg 21960  
actcctcttc ttcattggtg gcgcgtcgcg cctcggttcc ggctccgcgc tgcgccacgc 22020  
tgccgccac ccagctcacg atcagctcac ccagcggggg ctctgtcgcg aacagcagga 22080  
ccttctcccg gtagagctcg agcagcgcca ggaaccgcgc caccacctcg acggtgtgcc 22140  
cgagtcgca caccagctcc gagaacgtcg cgtgcctt ctccgccaac cgcaccgcga 22200  
gcagcgcgc gtgctcgcgc accgaaacgc cgtgctggtg gatgtggtcc agcgaaacgg 22260  
tcggcgggtg cttcggcccg aacaccgcgc ccgcgatctc ggcgaaacgc tgcggcgga 22320  
ccccgatcac cacttcgcgc agcaggttcg cgaaccgcgc ctgcaccgac accgaacgcg 22380  
gatagcgcg cagcgcgccg gcctccaact cgccgaacag cgcgccacc tgcctgtacg 22440  
cgcggtactg cagcagccgc gcgaagagca ggtcgcgggc ctccagcagc gccaggtccg 22500  
cctcgtctc gacctcgcgc gccggcagca gccgggcccgc cttcaggtct aacagcgtcg 22560  
cggcgaccac caggaactcg gtggttctgt tcaggtccca ccgctcgcgc agcgccttgg 22620  
tgtaggcgat gaactcgtcg gtgacctgt gcagcgtac ttcggtaacg tccagctggt 22680  
gctgcgagat cagctgcagc agcaggtcga aggggcctc gaagttgtcc agccgcaccg 22740  
tgaaccggcc gcctcggcg ggatcctcgc tcgactgcgc cgccgggtacc gcggggtctg 22800  
tactctccg ggctcctgca ggccgcgcac cagcaccgag tccgcgcgt tctgtcgag 22860  
gtcggcgagc accacggcaa ccgcttcgcgc caccacgcgc cctcggtcca cggcgatgcc 22920  
gtggtcggcc cgcaatgcca gccgcgcctg ctccaagccc agcaactcct cgtcggacac 22980  
gtacacggtg atcttcgaat cgtgcttgcg ccgtccagag ccgccccggc ccgcgcggg 23040  
tgccgcctgc tccgggtccg gcggattcgc gtgatccgtc cgaggaacag tagtacgga 23100  
caattcggcc gctcgggca gggcgaccgc acgcgtcctc gggcgatcac ctgcgcgcgc 23160  
agtgcgcggt acgcctcgc cccggccgag cggggcgccc agcgggtgat cggctcgcgc 23220  
gccaccgtg tctccgaaa ccgcaccgtg cggttgatca ccgcacgaa gacgatgtcg 23280  
ccgaacgcct cgacgacgc cgccatcacc tcgcgcgagt gcagcgtgcg cgggtcgaac 23340  
atggtcgcga ggatgcgcgt gatctccagc ctcggttca gccgtcccg cacttctcg 23400  
atggtgtcga tcagcagcgc caccgcgcgc aggtgaaga actcgcactc cagcgggatt 23460  
atcacggcgt ccgcggcggc caacgcgttc accgtcagca gcccagcga cggctggcag 23520  
tccaccagca cgtagtcta ctccgcagc gccggctgca gcaccgggtg cagcgtctgc 23580  
tcccgccca cctccgcgac cagctgcacc tcggccgcgc acaggtcgat gttgctgggc 23640  
agcaggtcca tgccctcgac ggtggtctgc cgcaccacgt catggatggc gaccggcctg 23700  
tccatgatca cgtttagat ggtctggtcg agctggtgcg gctggacccc gaggccgacc 23760  
gagagcgcgc cctgcggatc gaaatccacc agcagcacc ggccgcgta ttcgccagc 23820  
gctgcgcga ggttgatcgt cgacgtggtc tcccgacgc cgccttctg gttgcacatc 23880  
gcgagcaccg atgcgggacc gtgcccgtcc agcagcggcg gctcggggat gtgcccagct 23940  
ggccggccag tgggaccgag gccgcggggg ttcggcgcca ggccgaactc ctctcgcgcg 24000  
cccgcgtct cccggcgat gctcaggtcg accgcgccc gggaccgccc ttcggcggcg 24060  
ggctgcggca gcgacatggc gatctgactc ctttgggtgc cggtggggat caaacgcagc 24120  
ctaagtggca ttcgttacca gcggcaacgc gcctcgcgt ggtgtcggaa gttccctgtg 24180  
atctgcgtgt gatcaacgtc taccgcgtac accacccct cgcattcgca gcgtcaccaa 24240  
gtcactccgt agcgtgtgc cgcggtatgc tcgtcgcgta gacctccgc agcgtgttca 24300  
ccgtgatcaa ggtgtagacc tgggtcgtcg tgaccgaggc gtggccagc agttcctgca 24360

ccacgcggac atccgcg cg cctccacaa ggtgctggtg gaacgctgc cgcagcacgt 24420  
 gcggcgaaac cgagccgttg atcccggcgc gctgcgcggc ggtcttcaac gcgttccacg 24480  
 agctctgccg ggaaagccgg gtaccgcgcg aattcaggaa caccgcggca ctaccccgcc 24540  
 cgcgcgcggc caggccgggc ctggcccgtg ccaggtagcg gtccagcgcc gccagcgccg 24600  
 gtcggccgat gggcaccagc cgctggcggc cgcccttccc ttccagcagc acggtccggt 24660  
 tgggtgctgc gatgtcgtcg aggtccaacc cgaccgcttc ggagatccgc gctcccgctc 24720  
 aatacagcac ttccagcagc gcccggtccc gcaggccccg ggcatcctcg ccccgggcgt 24780  
 gatcgagcag cttcccagcg tcgccgatgg aaagcgcctt cggcagccgc ttcggcgggc 24840  
 tcggcggcgc gacatcgcg gcacgtcgga ccccgagcag cccctcggcg tcgcggaagc 24900  
 ggtgcaggcc gcggcgcc accaaagctc gggccgacga cgacggcgcc agcgcgggcc 24960  
 gctgcccggt gccttcccgc agctccgcca agaaccggga cagggtgctc gaacgcacct 25020  
 cgccgagccc agaaatccct gccgcgacca ggaattcggc ataccggcg aggtccctgg 25080  
 cgtaggagtc gagagtgtc cggcggttac cgcgctcgac ggcaagggtg tcgaggtacc 25140  
 cggatgatcg cccgcgcagg tccggcgga gctcctcgaa accctcacc gggcgcgaa 25200  
 gcgctggggc cggtcgggc attcggcac cgccgggga ggctccacc ggccctcgcg 25260  
 gaccgcgtgg gcggcgagca gcccggccac cccggggcg ttgacgatct ccccggccag 25320  
 cgccatccgc accgcctcg cgagcgggaa gcggcggtac -25360

<210> 4

<211> 45624

<212> DNA

<213> *Saccharopolyspora spinosa*

<400> 4

tcagccgacc ggcttccgc cgcgcagaat cgcagagccc atgtcgtggg catatttctc 60  
 ataatcgag accgcggcg cccagccggc gacagccggc ccgtacctgt ccgcgatccc 120  
 gtgctggtgc gcagcgagct cttcggcgaa ctgcggcatg aagtaccggg tctcgacga 180  
 cagctcctcc caatcgagga tctcgaaccc cgcgcagc agcgattcca gaagttgctc 240  
 agccaggcag atccgaaggc cggtcggcca cctgtccccg gacaccggca tcccgccgc 300  
 cgcttctcgt ttgacgacct cggtagcgcc gaggatgcca ccgggtttga gtactcgaag 360  
 gatttcccg atggcacggt ccggttcgga catctccaac agcgactgca tggcccaggc 420  
 ggcgtcgaaa gcattgtccg ggtacggcag ggacatggca tcgacgcacg agaagtccac 480  
 ccggtggctt agtcgcggt cgcgtgcgca atcagcggcg atggccactt gcacctggct 540  
 gacggtgatg ccggtgatct ggatcgcggt gtcgcgcgcg acgcgcagcg ctggttgctc 600  
 ggtaccgcac cccacatcga gcagtcgaac gccgccatcg agcacggtcc gttcggcgac 660  
 aaggtcgggt agccggtcgg cggcctgctg ccaggaagcc cgcccgctcg tctcccagta 720  
 gccgtggtgg atggcgagc ggccgcccgc gaccgagttc agcaacggcg tgaccaggtc 780  
 atacatctgc ccaacctgct gcgatgttg tgccgacct ggcaacaccg gtatgcctga 840  
 tctctgcaac gttcaccttc tcgtaatttt ctggaaccga gcaggaaaca cacagagaga 900  
 aaagacggca aactagccgt cagttaattc gcggttaccg ccgcagtcgg cgtaccgac 960  
 tataactate ~~cgccagggag gcaaacagcc agcgcaaat tccccatga atcccgagc~~ 1020  
 aaccaacggt gatcaccgta caggtccggc ccaccccgat cgagcggagc tcgcgatttt 1080  
 cactcgttcg cgtgaaacca ggtgcggcga ccaaagatc ttcgatggtc gcgtgacatt 1140  
 cctgccgtga cacactattg ttgccgcccg aagtcggtg gacgtcaacg gcgatcgaca 1200  
 tacgcagcgc attttccaaa tctggccaa tgggagtagg ggtgctgctg ctaatcgggg 1260  
 gaaagagact ggggtccgaa gtatgcgct actcgtcgtt cccttgcct atccgacgca 1320  
 tctcatggca atggtgcccg tgtgctgggc gctgcaagca tccgggcacg aggtcctgat 1380



cgccgcacca ccagagctgc aggcgaccgc gcatggtgca ggtctcacca cggccgggat 1440  
 ccgcggggaac gacaggaccg gcgatacggg tggaaccacg cagctgcgct ttcccaatcc 1500  
 ggcgttcggt cagcgcgaca ccgaggcagg ccggcaactg tgggagcaga ccgcgtccaa 1560  
 tgtcgcgcaa agctcgtcgc atcagctccc cgaatacctt cgactggccg aggcctggcg 1620  
 accgtcagtg ctgttggtcg acgtctgcgc gctgatcggc cgggtgctcg gcggattgct 1680  
 cgacctgccg gtcgtgctgc accgctgggg agtcgacccc accgcaggcc ccttcagcga 1740  
 tcgagcccac gagttgcttg acccggtgtg ccggcaccac ggactgaccg gcctgcccac 1800  
 tcccagctc atcctcgatc cctgtccgcc gagcctgcaa gcaagcgacg cgccgcaagg 1860  
 cgcaccggtc cagtacgtgc cgtacaacgg aagcggcgca ttcccgcat ggggcgcggc 1920  
 gcgcacctca gcacggcggg tctgcatctg catgggcccgc atggtgctga acgccaccgg 1980  
 gccggctccg ctgctgcgcg cagtagcggc tgccaccgag ttgcccggcg tcgaggccgt 2040  
 gatcgccgtt cccccgagc accgggcaact tctcaccgac ctaccgcaca acgcccggat 2100  
 cgccgaatcg gtcccgtcga acctgttctt gcgtacctgc gagctggtca tctgcgcggg 2160  
 cggctcggga acggcattca ccgcgaccgc actcggcatc ccgcaactcg tcttcccca 2220  
 gtacttcgac cagttcgact acgcgcgcaa cctcgccgct gccggggcgg gcatctgctt 2280  
 gccggatgag caggcccagt ccgaccacga acagttcacc gactcaatcg caacgggtgct 2340  
 cggcgacacc ggcttcgctg ctgcggcaat caaactcagc gacgagatca cggccatgcc-2400  
 ccatcccgcc gcgctggtgc ggacgctgga gaacactgcg gccatccgtg cctgatgaac 2460  
 tagtcacccg agaacagctg gatccggaga accgatgccc tcccagaacg cgctgtacct 2520  
 ggacctgctc aagaaggtag tcaccaacac gatttacagt gatcgggccg atccgaacgc 2580  
 ctggcaggac aacaccgact acaggcaggc cgctcgggcc aaaggcacgg actggccaac 2640  
 tgtcgcgcac acgatgatcg gtctggagcg gctggacaac ctccagcact gcgtggaagc 2700  
 cgtgctcgca gacggtgttc ccggggattt cgccgagacc ggtgtctggc ggggcggcgc 2760  
 atgcatcttc atgcgcgcgg ttctccaggc attcggagat accggacgta ccgtctgggt 2820  
 agtggattcc ttccagggaa tgccggaaag ctctgcgcaa gaccaccaag cggaccaggc 2880  
 tatggcgctg cagcagtaga acgacgtgct tggcgtagat cttgagaccg tccggcagaa 2940  
 cttcgcccgc tacgggctgc tcgacgaaca ggtcagggtt cttcccggct ggttccggga 3000  
 caccttgccc accgccccca tccaggaact cgccgtgcta cgactcgacg gcgacctcta 3060  
 cgaatccaca atggactcat tgcggaacct gtaccogaag ctctcgccgg gcggattcgt 3120  
 catcatcgac gactattttc tgccgtcctg ccaggacgcg gtgaaggggt tccgcgcgga 3180  
 actcgggatac acggaaccca tccacgacat cgacggcacg ggcgcctact ggcgccgcag 3240  
 ctggtgaacg gctcagccgt cctcgacgcc gagcgttgc cggggcacga accccggcgc 3300  
 ggcaggctca gcgttgagcc ctttctccac gaataccagg ttgtggtaga agtgcagggc 3360  
 cgccacgttc cgttccgtgt aggagggtc ggtcccgcac cgcgattcgc gtcctgata 3420  
 atgcaggccg tcgatcagtt ctttgagcat gtcgatcgag gtgcgctggg ccgcgggttc 3480  
 cccatcgca ccgcctgtag cgggccagta cgccgtctgg agatcctcga tgacgtacaa 3540  
 accacctggg cggacgtgcg gaaacaggga ttggaaggat ttcttgacgt ggtcgttgac 3600  
 atggctgccg tcgtcgatga caatgtcgaa cggcccgatc tttcccgcga tgtcttccag 3660  
 gaattccgca tcgtctggtt cacctcgag ctttcgcaact cggtgccctt cgttcccggc 3720  
 tttctgaaa atgtccagcc cgtaaacgag acctcgccgg aagtaccgct gccacatgcg 3780  
 cagcgaagca ccaccgagtt cgggtgcgtg ataaccaccg attcctatctt ccaacactcg 3840  
 eacccgggca tectggaate gggagaagtg gtgctcgtag tgttcgggtg accagtgcag 3900  
 gtccgcccac ttgtcggatc cgtagcggac tgccagcgca ccgaggcccg cgcgcctggt 3960  
 ggcggctgcg gccagcacgg cgtgcaccgc ggtggacacc ctgttccgga acgccagcag 4020  
 gcgctgcgcg ccggcgaggc cctggctcgg cgagaactgc gtcattgctgt ccgaccagcg 4080  
 gacttcacgg ctgttggtgc gcctgccgtc aaccgggccc aagagccctt ccagcagatc 4140  
 caccgcgtcg aaccggagaa cagcagctgc ttcggcgacc gccgcgagcc gcaggcctgc 4200  
 gtgatcgagc tcaacggttc tccggaccag ctgagcgcca gaagtgatct ccaggccgat 4260

tcgcaccgac tcgtccacg acagcggatc gcagcgtacg acgagcgt cgacgatggc 4320  
 gtccggccacc gcgtccagtc cggccacctg cactgcttcc tggagcttcg ccggggcccg 4380  
 acccgcccg agcagcaaac gtcaccac cgaccaggg gcaactgcga tctcactcat 4440  
 ggaaagtcat caccttttcg gttccagcgc atacggcacc atgcgcaccg cgatatcgat 4500  
 ccacaaagta gagccggcgg aggcgaccgg gtttcacgta tgagctgaaa cacacagata 4560  
 atcaccggga cgcgtagatg atctcggctg cgggcgaaca aagtggacca gtcagaaaag 4620  
 gaggggcggt gccgaattc catgaccgg caccatgaa tcgtcgaacc ccaggaacag 4680  
 agatcaccgt cgagcccgac gatcctcgtt atccggacct cgtcgtcggg cacaaccccc 4740  
 gtttcaccgg aaaacccgaa cgcattccaca tcgccagctc cgccgaagac gtcgtgcacg 4800  
 ccgtcgccga cgcgtgccc accggcaggg gggtaggggt ccgcagcggc gggcactgct 4860  
 tcgagaatct cgttgccggac ccggcgatcc gagtgtcgt cgacctctcc gagctcaacc 4920  
 gcgtgtacta cgacagcacg ccgggggcat tcgcgatcga ggccggcgcc gccctcgggc 4980  
 aggtgtaccg aacctgttc aagaactggg gcgtgacgat ccgaccggc gcatgtccc 5040  
 ggggtggcgc aggcgggcac atcctcggcg ggggatacgg ccgctgtcg cgcgattcg 5100  
 gttcggtcgt cgactacctt caaggcgtcg aggtcgtcgt ggtcgaccag gccggtgaag 5160  
 tgcacatcgt cgaggccgac cggaaactcca cgggcgcgg tcacgacttg tgggtggcgc 5220  
 acaccggtgg cgggtggcggc aacttcggga tcgtcaccag gttttggctc cgaacgcgg- 5280  
 acgtggtcag caccgacgcc gcagagctcc tgccacggcc gcccgcgaca gtgctgctcc 5340  
 gatcgttcca ctggccgtgg cacgaactga cagagcagtc attcgcctg ctctacaga 5400  
 acttcggcaa ttggtacgag cagcacagcg cgcctgaatc cacgcaactc gggttgttca 5460  
 gcacgctcgt ctgcgcacac cggcaagctg gctacgtcac gctgaacgtt cacctggacg 5520  
 gcacggatcc gaacgcggaa cgcaccctgg ccgaacacct gtcggcgatc aacgcccagg 5580  
 tcggcgtgac tccagccgaa gggctgcggg aaaccctgcc gtggttgoga tcgaccagg 5640  
 tggccggggc gatcgccgaa ggccggcgaac cgggcacgca accgaccaag gtcaaagccg 5700  
 cctacttgcg caccgggctg tccgaagctc aactagccac ggtttaccgg cggctgaccg 5760  
 tctacggata cgacaacct gcggcgccgc tgttgcgtgt cggttaccgg ggtatggcga 5820  
 atgccgtggc tccgtcggcc accgcactcg ctacgcgca ctcggttctc aaagcgtgt 5880  
 tcgtcacgaa ctggtcggag ccgcgcagg acgagcggca tctgacctg attcgcggtt 5940  
 tctaccgca gatgtacgcc gaaaccggcg gattccggt gccaggtacc cgtgtcgacg 6000  
 gtcctacat caactaccg gacaccgact tggccgatcc attgtggaac acctccggtg 6060  
 ttgcctggca cgacctgtac tacaagaca actaccgcg gctgcagcgg gccaaagcgc 6120  
 ggtgggatcc gcagaacatc ttccagcacg gcctgtcgat caaacgcgg gcacggcttt 6180  
 caccgggtca gccatgagga gtcctcacg atgtccaca cgcacgagat cgaacccgtg 6240  
 gaacgcatca tcttcgccc cggatccagt gcggcgagcc tggccgacct gaccaccgaa 6300  
 ctccgactcg ccaggatcg acccgtgctg atcgacgaga tctcttccg ccgggaaccg 6360  
 gccccgaca tcgaacggac cgaggtcgcg gtccagatca ccaccgagg cgagaccgtt 6420  
 gacttcgtcc tgacgtaca gtccggtgag ctgatcaagg ccgagcaacg accggtcgga 6480  
 gacgtcccgc tgcggatcgg ttacgagctc accgatctca tcgccagtt gttcggccca 6540  
 ggagctccca gggccgtcgg cgcgggagc accaacttcc tccgaaccac cacatccggt 6600  
 tcgatacccg gtccgtcgga actgtccgat ggcttcagg ccactctccg agtggtcgcc 6660  
 ggctgcgggc accgacgtcc cgacctcaac ttgctgcct ccactaccg caggacaag 6720  
~~tgggggggcc tgaactggtt caccgggata taagagcag acctcgggga gttccgtgat 6780~~  
 cgcccggtgc gcatcctgga gatcgtgtc ggtggctaca acttcgacg tggcggcggc 6840  
 gaatccctga agatgtggaa gcgtacttc caccgcggcc tcgtgttcg gatggacgtt 6900  
 ttcgacaagt ctttctcga ccagcagagg ctctgcaccg tccgcgcga ccagagcaag 6960  
 ccgaggagc tggccgcgt tgacgacaag tacggaccgt tcgacatcat catcgacgat 7020  
 ggcagccaca tcaacggaca cgtgcgcaca tccctggaaa cgctgttccc ccggttgccg 7080  
 agcgggtggc tatacgtgat cgaggatctg tggacgaact atgctcccgg attcggcggg 7140

caggcgcagt gcccggccgc acccggcacc acggtcagcc tgctcaagaa cctgttgaa 7200  
ggcggttcagc acgaggagca gccgcatgcg ggctcgtacg agccgagcta cctggaacgc 7260  
aatttggtcg gcctccacac ctaccacaac atcgcggttc tggagaaagg cgtcaacgcc 7320  
gaaggcggcg ttctgcttg ggtgccaagg agtctggacg acatattgca cctggccgac 7380  
gtgaacagcg cggaggacga gtgaacagca gaggggcgaa cacacaggca tttccgaccg 7440  
cggatcaggt ggagtccatc ttcgatgcgt tggcgccacgg gcgtcccctg caccacgggt 7500  
actgggcggg cgggtatcgg gaggatgccg gtgccacacc gtggtcggat gctgccgacc 7560  
aactgaccga cctgttcacg gacaaggccg cgctccgtcc cggagcgcac ctgttcgacc 7620  
tgggctgcgg caatgggcag cccgtagtcc gtgcggcatg cgccagcggc gttcagagtca 7680  
ccggaatcac cgtgaacgcc cagcatctcg ccgcgccac caggctcgcc aacgagaccg 7740  
gactggccgg cagtcttgag ttcgatctag tcgacggcg ccagctgcc taccgggacc 7800  
gtttctttca ggccgcatgg gcgatgcagt ccgtcgtgca gatcgtggac caggccgccc 7860  
cgatccgcga ggtccaccga atcctggaac ccggcgggccg gttcgtcctc ggagacatca 7920  
tactcgggt tcgactcccg gaagagtacg cggcggtttg gacgggcacg accgccata 7980  
ccttgaacag cttcacggcg ctggtcagcg aagccgggtt cgagattctc gaagtcaccg 8040  
acctcacggc acagaccagg tgcattggtc cctggtacgt cgacgagttg ctccggaac 8100  
tcgatgagct cgccggcgtc gagcctgcgg ctgctggcac ctaccagcaa cgctacttgg 8160  
gagacatcgc ggcgaagcac ggaccgggac cagcacagct gatcgcccg gttgcggaat 8220  
accggaacaa tccggattac gccagaaacg aggaagcat gggtttcatg ctctgcagg 8280  
ctcgaaagaa gcagtcctga tggcctccga gcacgccagc ctggtcggcg acgatctgcg 8340  
ggcaccgcg gatgatccct tctaccgacc gccgacgcg ctaccgcgg gtgtcccggg 8400  
cacgtcctc agggcccggc ccgtctcggc actgcgcggc acgggcgaac ccgtcgcagc 8460  
caaggcctgg caaatcctct accggtccaa ctccgccctt ggcattgccga acgccgtctc 8520  
cggcaccgtt ctggtgcga acatcccggt gccgcgcgaa gatcgcccca tcatcacttt 8580  
cgcagtgggc acccaccggc tcggtagcca agttgccccg tcgtacctgc ttcgaaccgg 8640  
aaccgagccg gagaccgagc tgatcgccgt ggccctcgac cgcgggtggg ccgtggtcat 8700  
caccgactac gaaggcctcg gtactcctgg aaccacacc tacaccgtcg gcaggcgca 8760  
gggacacgcc atgctcgatg ccgcccgcg tgcgcaacgg ctaccgggct ccggcctgac 8820  
gaccgactgc ccggtcggca tctggggcta tgcgcagggt gggcaagcgt cggccttcgc 8880  
cggcgaactg caccacact acgcacctga actgcgaatc cgcgtcgcgg ccgcagggtg 8940  
ggtgccgatc gatctgctgg acatcatcca ccgaaatgac ggggtgttca ccgggccggt 9000  
gctggccggc ctggtcgggc atgcgcgtgc ctacccgat ctgccattcg acgagcttct 9060  
caccgaagcg ggtcgtaccg ccgttgatca agtgcgcgag ctcggtgcac cggagctcgt 9120  
caccgccttc ctcgcccgcg agctgagcga cttctcgcac acttcgggc ttttcgagca 9180  
acctcgatgg cgcgcacgac tggccgaaag cgtcgcagggt aggaacgggt gcccggtggt 9240  
ccccacgctc gtctaccaca gcacggacga cgagatcgtt ccgttcgcat tcggcgagcg 9300  
actccgggac agtaccgcg cggcgggtac gccagtgcg tggcatccgc tctccgatt 9360  
ggctcacttt ccgcgcgcc tggccagctc gcgagtgttc gtctcgtggt tcgacgagca 9420  
cttctccgag ccgtccgcga tcagcgggtc gcgagatgcc aggtgagcgg atggcggtga 9480  
gcctctccag cggcttcccc gctggcttcc ggatgcccg tcaggcttcg aggcgaacta 9540  
ctaccgcgg atgccggcg gctatgtgga ccgcaccga cgggcctgac cgatgcctc 9600  
caccagcgcg gctgtgga teatgatte geeegaatet eeggeaeeeg eagggtctg 9660  
cacgcctgtc cctgctctga tgtcgcgtgc gaaggcgggt accgccttgc gaacctgatc 9720  
ttccgctggc aaggacaact cgtcgacaac gcccttcgc tcgattcgga tcacggcctg 9780  
ccactcggcg ggccggctga acgcccggtc gacgacgatt cccccacag 9840  
ctcgtacgcg ctgcggtagt ggtgcacgaa accgtatccg aggtgggcaa cgggtgccacc 9900  
ttccgattgg agcagcacgc tggccgacaa gtcgacgcc gactcctggg cctcgtgcga 9960  
gcttgcgccg agaaccgtga gcggaccgag gagaaagtgc cgagcggcac gggcgggata 10020

gacaccgatg tccagca g ctcgccacc gagttcgggtg cgatag gga tgtccgtgtc 10080  
gggaagcggc ggaattccga acacggcggt gaactcccg agctcaccga tctccccgga 10140  
ttgcagcagg tcgcggacca cgtcgtgccg gccgtgggtg aggaacagg aattctcccg 10200  
cagcagcagg ttcttctcc tggccagccc gaccaggcga gcggtgtcgg acgcgctgt 10260  
cgtcagcgggt ttctccgcaa gcacgtgttt gtctgcctca agcgccctgc cgtccactc 10320  
tgcattgatg ccaggcggca acggcacgta gacggcatcg atgtccggcc gctccaggag 10380  
ccgtggtaa ccagcaccg cctcgcatc gaatcgcgct gcgaaccgtt cggctttcgc 10440  
cggatcacgg ctcgccaccg ccaccacctc tgtttcggcc acgtcgcaca tcgcgggcag 10500  
catccgtgc cagcggaagg aagcgcaccc gagcacaccg atgcgcaccg gctttcgc 10560  
cgagctgggt atcgcacca cgcaccacaag ctatgcaggg aggcgaccaa gctgcgcgc 10620  
tggatgttca agaagtgggt gtcgccgagc agtcgccca actgtcccaa agtcatccac 10680  
cggaaagtgc tgggagggtc tgcgcgaag tctcatgca cctcgatgat ccggtacctg 10740  
ttctgcgcct ggtagaaccg accgccttct tcggacagga tcgattcgta ccgcacggtt 10800  
tcgggatcgg cggtagcac gtcgtccaag aacggcggcc agtcgttcga cggggtgctt 10860  
tggtagttgg ccacactgca ctggaccgtg ggagcgattt ccgcagtcca cttataccca 10920  
gcctccacc gagcgtggac caaacgtgc agcactccgc cgtccgctt gaccaacagt 10980  
gcgatctcac cttgttctc cggttcgatc atcggctgcg tccagctggc gacctcacga-11040  
ttggtcgcgg acaccgacac cgcgatcacc gcgaagtact tgcgctcctg gtgggcgac 11100  
tcggtatcgg tgcgatacca cttgtcgacc ctgctcagcg gaacgcgcgt tgcgcgaag 11160  
cggtagcggg ccttggtctc ctcgaaccaa ccgaccgct cggagatact cgacgaatcg 11220  
atgccgtgc agagcgacct ggccaccgcc tgcggaagg gtcctccga ggcggctagt 11280  
tcgggtccgg tggcggaatc gtggaacggg atgcaagaca gcaccgtccg ggtgtccatg 11340  
ttgacgatgt tgtcctgacg aaggagagcc agcacctggc cgagcgtcaa ccagcagaag 11400  
tctggcagga ctggcacttc ctctggact tccaccacca tgttcgggtt gcgcttccg 11460  
tagaaccagg aacctgttc agactggagc acgtctacca gcacgcggcc gcggccccgc 11520  
ccgaggaagt agtccacata ggcggaacg ctgccacggt gtgcctgcgt gtagttgtc 11580  
cgagttgcct gaaccgtcgg cgagagctgc aggacgttga cgttgcccg ttcattctt 11640  
gctgacatga ggcagtgcag cagccgtcgt atctccttga cgagaatgcc gaggatacct 11700  
acttcagcct ggttgatgat cggctggtgc cagctgggtg ccgcgccata gttggtctc 11760  
acctgcaggc cttctaccgt gaagaacct cgcgcgcat gaaccagggt ctcgctgctg 11820  
gcattgaatt tccatttcga caggcggtcg aacgggatgc gattggtctc gaagctgttc 11880  
tcgcccagcc gatcggcgag ccagcagtg aactccgtg tcggaatcat gccattgcaa 11940  
gcgctcagcg cagagtcgac gaaccgcgc gtgtgtgtgc tgcgagcgg cgacgagca 12000  
cttgctcag cttcgacaga actgctcatg ccaatccct cactgccgtg acgatgtgca 12060  
cgccggctcc cagggcgctc gcggtctggt ccagcggcct ggcgaggatc atcgaagcga 12120  
tccgcccgtg aggaccgcga gcacctcgtc ctggcgcttc gtcattgaaga aatgcccgcc 12180  
ggagaagaca cggacgtcgg ccacggcctc ggtgtgctct cgcaggctt ccgctcctc 12240  
caaggtgacc ttcggctcgg catctaccac cagcacggtg atcgggcaac ccacctctc 12300  
tcttgccgca cagcggatg tctcgaccgc acggtagtcg ttgctcagcg ggcgccacga 12360  
tctcgtcgg aaatcgggtg gagcgcaaac ctctggggag atcaaacacc cattggacga 12420  
ttctggacaa gaatctgact tcgggactcg cctggcatcg tcccggtcac ggcggaatcc 12480  
gagtcagcgc aacagcgatg aacacctgcc aaaccccgcc gtccatgttt ccggccacac 12540  
gttgctcacg gatggcctc agactgccca gtgcggcggt tctcacgttc ggcaactagc 12600  
ttttccagac gtggaacgat gtcgtgcggg ctgggacttt cctcgatttc acggcggtc 12660  
agcgcgcgt tggcagcga cagcggctc tgcagcaggc gggccagctg acctcgcacg 12720  
tcgtcttcgg taaacgtcgc ggggtcagg accagaccg ctccccgatc ggctaggagc 12780  
tccgcccac gagattcgtc ccagaaggc ccaggagaa tcaactgcg tacgccgttg 12840  
accgtggcg tttcctgcgt cgtcgtcgag ccatgatgga tgatcactga acacgactcc 12900

agcagttcgt tgagcgggtac gtatttcgtgg acccggacgt tcgagggcaa ctcccccatc 12960  
 tcccgacttt cgtcgtcaga caaggtggcg atcacctcga cgtcgcgcct ggccgcgcgc 13020  
 cgcaacaacc tttccaccat tgcttggtcc tgggtacttt gctccccgta ctgttcggtc 13080  
 accctgctca gccgcgcgtt ggtcagcccg cgcgtgatgc agacgcgcgc cttcgtcggc 13140  
 cgttcgcgca accactcccg caacaccgcc ggaccgttgt acggcacgaa acgcatcgag 13200  
 atgtagtcca agtccacagg caggcgcac caggatggaa tcggatctat ggtcgttg 13260  
 cccgtcacga tctcttcac gaacgtggct ccgaacttgg cgagcttcgc tcccagccac 13320  
 gtcccgagcg ggtcgcgcgc ctgctcaggc ggcttcgatt cctgggtattc gaggaaccgc 13380  
 gaccgcagcc accccgacac atcgagggcg acgagcatcc gcacgtgtcg tgcgcgcgagc 13440  
 gctcgcgcga caactggccc cgaacacacc atagcgtccc acacaacgag atcaggctgc 13500  
 catttctcgg cgaaccccat gagatcgtcc agcgtcggc catccacaag gtggagtgc 13560  
 tccacggcgt ccgtgtctct gccggagttg attgacagca gctcgtcgaa gagttccgga 13620  
 tgccgcccct tctcgtcgaa cgagaccccc ctgccgagaa cgagttcgtt cttggccgcc 13680  
 aaggagatga ggtcgcgcgc gtcgccgacg ggaaccgcgc tgagtccctgc tccggtgacc 13740  
 atcgacacca tattcgggca gatggcgaca cggacctcgt gcccgcgcgc acgcaaccgc 13800  
 cagcccaacg gcaccaggtt gaagaagtgc gaactcgcgc gcagcggggg gaacaggaca 13860  
 cgcgtgaact ctccgatcgc aattgaacac ccgggaagca tgccaagaat cacagaaatc 13920  
 tctgatatcc cccgggaaac gccgctttcg caagccaaat cttaggcctt ccagggtgatg 13980  
 gtagcgatct tgacaagcgc gagcaggtcg ttcccgtag cctgggctct accgagtcgc 14040  
 gtgtgccggg tagatcgagg atttctgagt caatgagcgc ttctccttgc tccgctgtcc 14100  
 tgatgtcccg caccgcacgc aaccagggca ggaaggtgta aggcgcgcgc acagcacact 14160  
 gtcccgctgg gacgtcataa cgcgattcgc caggggcac cgtcatctcc tgaaggcaag 14220  
 gcgcgaagac tgatcgtcgc ctgcatgcag agccggaaaa ccagagcgcct ggggaaagg 14280  
 cgcgcagag tgacttcgtg tgacgacact tgcgctaccg ctactgagat ggtgccgat 14340  
 gccaaaggac ggatattggc atccgtacgc gattaccacc gcgaacagga atccccgacc 14400  
 ttcgtggctg gatcgacgc gatccggcca tcgggcgcgc tgctcgacga ggacgaccgc 14460  
 gtggcactgg tggaagcgc gctggagctc cggatcgcgc cgggcgggaa tgcacggcga 14520  
 ttcgagagcg agttcgcgc cttcttcggc ctccgcaagg ctcatctcgt caactccgt 14580  
 tcgtcggcca atctcctggc actgagttcg cttacctccc ccaaactcgc cgaggcacga 14640  
 ctgcggcccg gcgacgaagt gatcactgcg gcggtcggct tccccacgac gatcaatccg 14700  
 gcggtccaaa acggactcgt cccggtattc gtcgacgtgg aactgggcac ctacaacgca 14760  
 acgccagacc gcatcaaggc cgcgctcacg gaacggacgc gagccatcat gctggcgcac 14820  
 accctgggca accccttcgc cgtcgacgaa atcgcggaga tcgcaaaaga acacgagctg 14880  
 ttctcgtcgc aagacaactg tgatgcggtg ggatccacct accggggacg gctgaccgga 14940  
 accttcggcg acctgacaac ggtcagcttc tatcctgccc atcacatcac cagcggcgag 15000  
 ggtggctcgc tgttgaccgg cagcctggaa ttggctcgca tcatcgagtc gctgcgtgac 15060  
 tggggacggg attgctggtg cgagcccggc gtggacaaca cctgccgcaa gaggttcgac 15120  
 taccacctcg gtacccttcc accgggctac gaccacaagt acacgttctc ccacgtcggc 15180  
 tacaacctca agaccaccga cctgcaggcc gcaactgcgc tgagccagtt gagcaagatt 15240  
 tccgcattcg ggtcggcacg ccgcccgaac tggcgacggt tgcgcgaagg gctgtccggg 15300  
 ttgccgggccc tgctgctgcc ggtagccaca ccgcacagcg acccgagctg gttcgggttt 15360  
 gcgatcaca teagtgaggc cgcgggttc accggtgcgc cctggtgaa cttctggaa 15420  
 tcccgcaaca tcggcaccgc actgctgttc ggcggtaca tcaccggca cccggccttc 15480  
 gagcaggtgc ggtaccgat ccgcgacgc ctcaccaaca gcgacatcgt caccgaccga 15540  
 accttctggg tcggcgtcta cccaggcata acggacaaa tgatcgacta cgtcgtcgaa 15600  
 tcaatcgtg aattcgtggc caagagttcc tagcatccag catggcggca tctcggagga 15660  
 tttcagcaac gtgatcaacc tgcaccagcc gatcctcggc accgaagaac tcgacgcgat 15720  
 cgcggaggtg ttcgcctcca actggatcgc gctcgggccc cgcacccgga cgttcgaggc 15780

cgaattcgcc caccacgagg gagtggatcc cgaacaggtc gtgttca actcggggac 15840  
tgccgcgctg ttccttaccg tgcagggtgct cgacctcggc ccaggcgacg acgtgggtact 15900  
tccttcgata agcttcgtgg cggcgggccaa cgccatcgca tcctccggtg cccgcccgtt 15960  
gttctgcgac gtcgaccccc ggacgttgaa cccacgctg gatgatgtgg cgaggggccat 16020  
cacgccggcg accaaggccg tattgtgtct ccactatgga ggatcgccgg gagaagtcac 16080  
cgcgatcgcc gatttctgcc gtgaaaaggg cctcatgttc atcgaggact ccgcctgcgc 16140  
ggtggcatcg tccgtgcacg gcaccgcttg cggaaccttt ggtgacctgg ccacgtggag 16200  
tttcgatcg atgaagatcc tggtcaccgg ggatgggggc atgttctacg cggcggtacc 16260  
ggagctggcg caccgcgcaa gacgactcgc ctaccacggg cttgagcaga tgagcggatt 16320  
cgattcggcc aagtcttcca accgctgggt ggatattcgc gtcgaagaca tcggccagcg 16380  
gctgatcggg aacgacatga cggcagcgct tggcagcgct cagctgcgca aactgccaga 16440  
attcatcaac aggcgtagag aaatcgctac gcagtagcag cggttgcttt ccgatgtgcc 16500  
gggtgtcctc ctaccgccga cgctaccgga tgggcacgtc tcgtcacact acttctactg 16560  
ggtccagctg gctccggaga tccgcgacca ggtggcgag caaatgctgg aacgcggcat 16620  
ctacacgagc taccgtacc cgccctgca caaggctccc atctaccgcg cggactgcaa 16680  
gctgccttct gcggagcag cctgccgcag aacactcctg ctaccactgc acccaagcct 16740  
tgacgacgcc gaggtgcgca cgggtggtga cgagttccag aaggccgtcg aacaccacat 16800  
cagccaaaga tcaccactcc gaaagtgagg atgtcgcgcg tgagcgacac attcgcagaa 16860  
acctcctcgg tatacagccc agatcatgcc gacatctacg acgcgatcca ctccgcgct 16920  
ggccgggact gggcagccga ggccggggaa gtagtccagc tcgtacgcac caggctgcc 16980  
gaagcacagt ccctactcga cgtcgcctgt gggaccgggg cgcacctaga gcgattccgt 17040  
gccgaatacg cgaaggctcg ggggcttgaa ctgtccgatg cgatgcggga gatcgcgatc 17100  
agacgagtc ctgaggtacc gattcacatc ggtgacatcc gcgatttcga cctcggcgag 17160  
ccattcgacg tcatcacctg cctgtgcttt accgcggctt acatgcggac cgttgacgac 17220  
ctgcgacgcg tgacgcggaa catggcccgg cacctggccc ctggtggagt cgcggtcatc 17280  
gaaccctggt ggtttcccga caagttcatc gacgggttcg tcaccggagc cgtcgcgcac 17340  
cacggcgagc ggggtgatcag ccggtatcgc cactcgggtc tggagggccg tacgagccgg 17400  
atgaccgtcc gctacacagt cgcgaacccc accgggatcc gggacttcac agagttcgaa 17460  
atcctctcgc tgttcaactga ggacgagtac accgcgcgc tcgaagacgc agggatccgc 17520  
gcggaatacc ttcctggagc accgaacggc cgaggcctgt tcgtcggaat ccgcaactga 17580  
gcccggaaaca acgacgcaag acctggctgg gcaggccccg ggactcactc gaccgataag 17640  
ccgacgtgac ccgggatcac ttcgatcgtc tctgatgcgg gcaagttttg accgtcgtga 17700  
tctaggtgcc aaccaggtcg acgaagccct gggccctagc cgcgatcgtc ggcaaggctg 17760  
ccgaggacga tcccgcctcg aatctgcgac cactcgcgc cagctgttcc ggacagctt 17820  
gtgatccgag ccgacgagaa ccactcgggt cagcccgga cctgcggctg ggaggtgttc 17880  
ggctcctgag cagcccaaaa gcaccggcca tcggcgaaaga actcgaccgg ctcccgcgc 17940  
ggatgcgcct acccgacctc cgcaaagccg caccgcctat gctggcaacc gccacgcct 18000  
gacctaaagg gccggtgagc cagggctggc ccagtgtcgc caggcaaccg cggcgacccg 18060  
gcggtaacgg cgcgcgaagt acaggtgtc aatgaccgag acggcgaaagg ccgacatccg 18120  
taccacaggg aacaggtgtt gttcggcgta cccggtcgca gcgagcctct ccgacttcga 18180  
ttgcaccatc gccgactgag cgcctcacca gggcacattc ccggcgctgt cgaacatgcc 18240  
tccggtcggg ccgtcgtctg gcagcgtggc gagccgaatg gcgatcctgg caccgtctgc 18300  
cggcgtgctg gttccgtgga agccgttaag gtcggtcgcg acgtagccgg ggcaggcgtt 18360  
gttgattttg atgttggtat cgctgagttc cttggcgtag tggatggtga tcgcgttgag 18420  
gaacgtcttc gacggtgagt aggtccgct gatcccgccg aggtcgacgc ccggcgtggt 18480  
ttgcaaggtc agggaaagcga cgtggctgga ctggttgacg atccgcgggc gctcggagcg 18540  
gcgtagcaac ggcagcatag cgttggtaac ccgaacgacg ccgatcacgt tggctccac 18600  
caccgcccgg aggtcgcgcg gtgtgacggt cgagggtcc tccggccatg ccccgcgat 18660

gcccggcggtta ttaaccagca catcgagccg gccggcgcgcc tccctcgagca gagccgcagc 18720  
agccgcgacg ctgcgctcgt ctgtcacgtc cagggatacc gcgaacgcat cgacgccgtc 18780  
cgacgcgaat ttgccacgg catcctcccc gcgctggtgg tcccggtgcc cgattccgac 18840  
gctccacccc agcgcgccga gcccggccgc gatctcgtat ccgattccct tgtttgccc 18900  
ggtgaccagt gcaatcgtct gttcgtctcat gctttctagc gtggcccga accggcgggg 18960  
aaccaacacc atctaggtct cacaccgata ccagctgat atgaatcagc ggtaggctcg 19020  
acggatggag acgcgggagt tgcggtactt cgttgcatc gccgaggagt tgcacttcgg 19080  
ccgggccgcc cagcgcttg gcacgcccc gccgcgctg tcgcggacga tcgccagct 19140  
cgagcaacga ctcgagatcg tgttgctgca acgcaccagc cgcaaagtct cgctcaccga 19200  
agccggggca atgctgctga ccgaaggccg ggcgctcctc ggcgcgctgg cagcagccga 19260  
gcgacgcacc cagcgtgccg cgacgagcca gccctcgcta gtccctggctg ccaaggccgg 19320  
cgccctccggg gagctgctgg cgaagttgct cgacgcgtac gccgccgagc cgggagccgt 19380  
ggccgtcgac ctgctgctct gcgaatcccc gccccagaaa acgctgcatg acggccggggc 19440  
cgacgtggcg ctgttgcatc aacccttcga cccgacggcc gaactcgaca tcgaaattct 19500  
gaacaccgag caacaagtcg ccattcttcc gacctcgcat ccgcttgcca gcgagcccca 19560  
tgtacggatg gcggatgtca gctcaactgcc ggatctccc cttgcgcgct ggccccggccc 19620  
cgacggcgtc tatccagatg gccccggcgt ggaagtacgc aaccagacgc aactgttcca 19680  
aatgatcgca ctcgcccgca ctaccgtgg catgcccga tccagtcgag tcaacctgct 19740  
cgaaggcctc gccgccgtac cggttctaga cgcgccggac gtgacgacag tcatcgctg 19800  
gccgccccac agccgctccc gagcactcgc cggcttggtc cgcgtggcca cactcctcta 19860  
aatcttcgct cctcaattg cccgcagcta ggacgaatgc cgtgcgactg ccgaaaacga 19920  
ccgcgcgcaa cacctcatgt gcctcagcgg gagaagacgc tatgcggatg ccgcagagac 19980  
ctcacggccg tcccgctgcc aggagccagc ggtcccgttc gagctggtag gtttcggcga 20040  
tctgctccgt gacgcggtcg atgtagccga agctggttgc gagcattcgt tgcgtggtcg 20100  
ctgcggatac ggccctcgtc gcgcattggc ggcggtatct gtcgaggcat cactacagga 20160  
agcggaaagt cccgatgcgg ttggccctga tcagcgcggt gatggaaatg ttctgccagt 20220  
cgctcgcgat actccaacac cgcagcagga cccggcgaat atcccaacgg aatctcgaat 20280  
tcgagaacgt gcggcaatgt ggcatatatt tcctcaacac tcgtccgtag aatgctgagg 20340  
acgatctcat cggctcgcaa ctgcggagcc tcggtaatca ggtgcgtgta gacgtctcgg 20400  
ctgacttcgg agacgcacga cgcgacgtc gatgccacgg ccgcctgtag cttcgaggaa 20460  
tcgaagctgt ccgtggccat gacacgacgg tactgcggcg ggagcagaag cagcgcttcc 20520  
tgcccgcatg gcgcgcctgg agaagatcag caccttcggc ctgaccgagc cgggcgcacg 20580  
gcaccgacgt ggtccagctg gagctcggag acacagcgtg gcggtgttca gcgcgcag 20640  
gtcccgagc cagcgctttg gcgtgcggcg cttcgggaga gctgggcat tacaggttgt 20700  
gcgagagcag ctggctcagc acctcgtcga cgttcatcca ccggtgttca tcaccaggtt 20760  
cgaccacgtc gtaggtgtcg ttcaagaagt ccgcaagctg gtcagcattc acctcgaagg 20820  
aggcctgccc cgacggcgag ctcatctcga tcacgaccaa ccccgaaaaa ccccgtcgag 20880  
ggccgatccg cacatcgctt tcgctgctt cggcgatcag cccgctcgcc agcaggtcgc 20940  
ggcgcatcac ccagtcacc tgaccgtccg tgccatcgtt gagcttcatg gagatctcgt 21000  
acggattgcg gctgtcgtac cgcagctcca cgcgactgg cgcgatcaca cccgccgggg 21060  
ctaccaggtt gaagatcatg gtggcgagca ccgtgccgtg atcgtttgcc atgatgtcca 21120  
cgctcccaag ccttggttcc caataccttc ataaagaagg atgccccacc cgggttcata 21180  
ccttcttcac gaagtcttca gaccctgggt gcagtccccc tcgggctccc agcgaggaac 21240  
cttcaactcac actgaacca cgggactctg ggcccggaac tcccgaagct gacccaaggt 21300  
gcaaagcaac cggcgagtag cttttcgtc cgcagtcga tcaatgacgg ctccgggaat 21360  
gtctcgtgac catgctaggg cgtactcgc gaaaaatgg atggatacgg gctgaccagc 21420  
aggaacgcgg tcagacggtc tggctcctcg gaaagaatcg ggcggtttg tctggcgcg 21480  
gcgtcaggac tggccaagg cccgttcgat cgcgcggcg gcccgggccg cggccgccc 21540

gtcgtgctgg agctgcaagg ccgcgttcag tgccagcccg gtggcgagga tctcgaacag 21600  
 cgcttggtcg atgtcgaacc cggcgggag ctcgcggttc tccaccgccc cggtcagggtc 21660  
 cgcccgagc tgtccccgcc agcgcgacca cacctcggcc accgcgtcgc ggaccggcc 21720  
 gggcgggccg tctactcgg tgagcgccgc ggtcatcagg cagccgcccgc gcagcagtgg 21780  
 cgcttctagg tagccacagg cgttggcgca caccgcgcgc agccgcccga gaccggcgg 21840  
 ctccggcagt gcgggctcga ccaccgggtg ccagaagtcc acgaatgcct tgtccagcgt 21900  
 ggagatctgc agtgtctcct tggtgccgaa gtgcttggtc accccggact tgctcatctc 21960  
 cagttcctcg gcgagccggc cgatggtgat gccgtccagc ctttcctcgg atgcgatctc 22020  
 ggcggcacgg tcgaggatcc ggctcctggt ggcctgtgct tcggccgctg agcgtcgcgg 22080  
 tgacatgtga cgaggatagc gtacgcgcgt tcgctattga tttagtgaac gtgcgtacgc 22140  
 taaattccgc ttcaccacag gaccgaagga gtgcgagatg cgggtgagac gactgggctg 22200  
 ggccggactg gaaatcgagg cgggcggcaa gcgactggtg atcgactacg tacgggacct 22260  
 ctcaccgctg ttcacggcgt ggaaaccggc cgagaggctg gcggtgccga gcgggacggt 22320  
 caccgcccga ctggtcacc acctgcaccg ggaccacacc gacgcggccg cgctcgcggc 22380  
 cgcgctgaca ccgggggac cgggtgctccg accgcgccc ggccacggcg acgacgtgga 22440  
 caacgtgacg aactgcggc cgagcgcgga gctgacactg caccgactgg ctgccagggt 22500  
 cgtggatgcc tggtcacccc gcgacctcg gccgttcgc gtcaccgagg tccccgcgt 22560  
 cgacgggctg ggcgaccgc agctgaactg ggtggtgcag gccgacgggc agcgggtctt 22620  
 ccacggcggc gacacgatgt tccacggcta ctggtggctc atcgcgcgcc ggttcagccc 22680  
 gttcgacgcg gtgttcctgc ccgcaacgg cgcggtggtc gacgcgccac acctccagcc 22740  
 gccgagcccg ctgcccgcgc cgatggaccc gaggcaggcc gccgcagccg cagagatcct 22800  
 cgacgcccgc tacgcagtgc cgatgcacta cgaggcgga cagccggaca agatcgcggg 22860  
 ctacgtcgag gtctccgacc cggagaatga gttccgcagc cagccggac accgcgcaca 22920  
 cgtgctgccc atcgaggaat ggctggacct ggccacttga ggcggaact ccgcgcgcgc 22980  
 ggcgggctca ggtcgcaaca gaacctaggt ccgatggatg gtccggaatc gcgcgggaat 23040  
 cggttgctcc aagacactct ccggtcgcgc cggcgtcaat cacaagattc agcatccatt 23100  
 actgctgga accacgatca aatgcgacac tccgggttga gccgaaccca cgggtccagt 23160  
 cctggtcctg ggcgggacgc accttggtt tccgcgcgag gaacctcac ttgcgtgaa 23220  
 cccacggcgc ccgaagcgtc ccggtagcag acgcccga aa ccagcgctta gctctggaac 23280  
 gctggctgcg gaccgggccc agaggcgagc gctcgggtgg tgctgagttc ggggacgtac 23340  
 ttgtagatcg tcgagcgga gacgcaccgc acgcttgagc ttgaacggcg cagcacgcga 23400  
 cccgcccag ataccggtta gcgtggtcag atcaacactg cccgttggtc accattagcc 23460  
 gcataatgccc ggttaccatc gattcgatgc ggatccggtc cgtgtcaaag ctctctacgt 23520  
 cgacgggtac ctgaatgccg cgtcaacga cgcgtcatcc gggtcagccc ttctcgagca 23580  
 ggcgagcgcg ctggcgggcg agttgggcga ccaatcaggg gcggcctacg tcgcgcagat 23640  
 ctccgggctc accgcctgt tcaacgaaga tccggcgcg gcgagcgcg tcttcgaaga 23700  
 cgctctcgcc aagcatcggc tgattggcga tcacgtgccc gccgcctacg atcagatcga 23760  
 gctcgcgttg gcagtcgcgt tccttggtga tcatgagcgc gccgctgccc tgttcgcgg 23820  
 gtgccctgcc gtcaaggaga cctacagcga acagtggatg agatcgcttg ccctgtgggc 23880  
 gaagggaata gcagacctac tagcgggcaa ttaccggcaa gccaatgctg cggaactgga 23940  
 gagtcttcgc ctgcggctcg gcttccatga gcagtttcag atcgcgttgt gcgttgagat 24000  
~~tetggcctgc atagecagcg cagaagggtga cccggcaagg gctggagact tgttcggcgt 24060~~  
 gtcccaaacg atcaaaaaga acgtcgatgc gtcgctcgcc ggccacaaac atgttgccgcg 24120  
 cctacacgac cactacgagg cagtagcacg tagctcactc ggcgacgaga cgtttcagaa 24180  
 aggcgttcac caccgacac agctcgactt cgacgaggcc atcgcatgga tacggaatcc 24240  
 gggcaaagcc ggggaaccag caagctccag gacggaagaa acgcacgcgt cgggtgcttac 24300  
 tccgcgcgag cgagaggtag cagcgtggt cgcacagggc accaccaaca aagaaatcgc 24360  
 cgccatcatg gtgatctccc gccgcaccgc tgaggccac gtcgagcaca tctgaccaa 24420



actcggttc	accactcggg	cgcaaatoeg	gacctgggtc	attcaacaga	agaccgcgcg	24480
acgatgcgac	tgacaaccgc	cggttttcgg	attggcgtgg	accagtagaa	gctgccccgt	24540
ctgcgagacg	gccccctctc	acagcctctt	ctccacagtc	gatcaacgac	cgaacggggac	24600
cgcttgagga	cggcacccga	tgcccaccgc	gcgagagagga	gcacttggtc	atgcgcacac	24660
tgatcgagaa	cgctgccgtg	gtgacgatgg	acgacgttct	cggcgacttc	gaacgagccg	24720
acatcctcgt	cgacaacggt	gtcatcgctg	agatcggggc	gaacctcaac	gttggcgaca	24780
ccgagcgaat	cgacgcctcg	tcgatgatcg	ccatgccggg	catggtcgac	acccatcgac	24840
acacctggca	aaccggggtc	cggggcattc	tggcgagacg	gaacatcctc	gactacctga	24900
ggggattccg	gctgcagatg	gcgacgaagt	accgcccgcg	ggacatgtat	gcggggcaact	24960
acctcgggtg	cctcgattgc	cttaactcgg	gggtgaccac	gctggtcgat	tactgccaca	25020
acatcgtcac	cggcgagcac	gcgcacgccg	cggtcgctgg	actccgcgac	gcgggcgtgc	25080
gcgcgctgta	cggccacggc	ctcctgcccc	tcacgtcgaa	cacgtggtcc	gagaccaaag	25140
gcgggctgga	cgagtcgggtg	gaagaaggcg	acttcgcccc	gcgggcgcgg	ctggccccggg	25200
aaatccgcgc	ccagtacttc	ccgagcgagc	aacaattgct	gcgcttcggc	atcgccccgc	25260
aggaactagc	catcgctccc	ttcggggacg	tcaagcagga	attcgagctg	gcccgcgagc	25320
tcggcgacag	catcacgttc	cactgcaacc	aagtcacgcg	gcgcaacctg	ttcaaggaca	25380
tcgaggcgct	gcactcccac	aacatgctcg	gcagcgacct	tctcttggtg	cacgggacat	25440
tcagcactga	gaacgaatgg	cacctgctgc	ggggtaccgg	cacctgatc	tcggtttgcg	25500
ccgagaccga	gatgcagatg	ggcatgggat	tcccggtcac	tcgtgaggcc	accgagaaca	25560
cgccggggcc	aagcctcggc	atcgactgca	cgagcagtac	ctgcggcgac	atgatttcgc	25620
acgcacggct	cgctctacag	gtgacccgat	ggagggacga	tcaagaagac	tacgcgcggt	25680
cgacactgcc	cacggtcatg	cggtggaaaa	cgcgcgacgc	cttacgggtg	ctcacggtca	25740
acggtgcccc	cgccgcgggc	gtcgacgacg	tgacaggaac	cctggcaccg	ggcaagcggg	25800
ccgacatcgt	attgctggac	atgtcaggca	tcagccaggc	cggttggaac	cgccacgacc	25860
cttgtggcgc	gatcattgcg	cagaccaact	cgggcaacgt	gcacacgggtg	ctggtcgatg	25920
ggcgggtggt	caagcggaac	ggccggctgg	tgacagtcga	cgtcaacgga	gcgcttgcaa	25980
cgctcgcgga	gtcgacgggt	catctctacg	accaaattgg	ccagcacggc	ggattcatcc	26040
cgcaacctcc	agccgagcta	ccggtgttca	accgctgaga	accaagccag	ctgtggctgg	26100
cgcgccattc	gctgcgtgag	gtgacccggc	tggccggggt	cgaccggaaa	accgtcagac	26160
gctgctgtgt	gacctgcacg	gacgagggtta	tcggcatagg	caaggccagt	cgacgacgca	26220
cggtcgtctc	tcatgcccgc	cccaccgtca	gcccgtgctg	agcgccgatt	cgagcaacca	26280
cgcccgggcc	cctggaatag	ccgcttccac	ctgcacggag	caggatttcg	cctcgtggac	26340
cgatgatcgt	ccccgtgaac	gcgtgggggt	gcccgtgcggg	aacgcgcggt	cttcgggggg	26400
gaaggcgggc	tgccgggaacc	ggcgttgact	gacctcgt	gcgcattcga	gcacgacggg	26460
ccacatcaat	ggggagggcc	cggtgacgac	aaagaccgtc	aggtgtggtg	gcaggacgcg	26520
gtccgaacgc	accgtgaccg	ttgaacggtc	ggccgagctg	tcgtttcttg	ccgcagcaac	26580
agcaactcga	cgaacgcgcc	ggcgatcttg	ccgcgcgggg	gccaccaacc	gcgaactcat	26640
gacccggtc	aacacatcaa	cccagccccg	ataacagggt	tcgagccctg	cctgtcgact	26700
cgtaggcaatc	gccctaaagg	agggccagct	gtggttacgg	atcgtgttcg	cgtgctcctt	26760
tgtgaagcgg	gcgtgagtga	gttcataagag	cccgcgcgcg	gttcgtgagt	aagttcttga	26820
acatggagaa	cgccggaccc	agccggacag	cgctcgctac	cgcgtagccc	cgcgccctatc	26880
accagatcgc	tgaccgacca	cagatcttca	ccgacccctt	ggcggtacgt	ctgctcgggc	26940
tcaccgcaga	ggaactgacc	aaattggcca	cgcccacaac	cgatcacctc	ggcgccgggg	27000
caatcgatca	gccccgcgcg	ctgttcttcg	ccgcccgtgc	ccgcttcgcc	gaggacgccg	27060
tggccgcggc	catcgccgat	ggtgtacggc	aggtcgtgat	cctcggcgcg	ggcctggaca	27120
ccttcgccta	tcgcaatccg	cgcccggaact	tgccgctctt	cgaggtcgac	caccccgcca	27180
cccaagcatg	gaaacgcgaa	cgcccttgcta	gcgcagggat	cgaccgccct	gagacgttga	27240
ccttcgtacc	ggtcgatttc	gaaacccagc	cgctggcaac	tggattggaa	tccgccggat	27300

tcaaacggac ggacccggtg gtattcgtgt ggctcggcgt tgtcttcat ctgaccgcga 27360  
acgccgcoca cgccaccctc gaatacatcg ctggtcaagc ccagccggtc gaggtggtct 27420  
tcgactacct gcagcctgcg gccaccgacg aggaccgtgc gcaccagcgg gcacgtgccg 27480  
atcggctggc cggcgccggc gaacccttgt tcagctactt cacaaccgac gacatcgccg 27540  
cacagcttca cgccctcggg ttcaccggca tcgaagacca ctccgccccg gacctcatcg 27600  
ccggctacct ccacggatcc gcaggtttcg agggcgaacc accccgcgca ctgcgcgca 27660  
gccgcatact gcacgcgagc cgctgaggtt tgccctgccg gcctacttgg accacatctg 27720  
tcgcgacggc tagcatctcc cacatcgctc tctgaactgc acaaactcaa aaaatcgcg 27780  
gagcgcacgc gggtagggcc ccggcgaaca agccctagcg atacatcact ttgaacaagg 27840  
tcgcgtctcg ataggctgcc aatgcgcgct tgcggtttcg gcgaaccgcg cggctcgactt 27900  
ggctcgtggc tagtcaagga accgctgcag ctcatcgcg gtgaacgggc ggtcgtcccg 27960  
atcgccctca taggactgga gatgggcgag tgagttccac tcgtgcacat tcgccaccgg 28020  
gtactcgtcc ggcccgaacg acgctcgca cgccagcgcc cagccatacc ggctgtcaca 28080  
gaggaattcg ttgaacagcc tgagactgcc ttggtaggcg cggattgtgg atggcgcgag 28140  
atgcggttct ccggtcgggtg attgcgacca ttcgtccaca tgcgacgggtg tccaccgcca 28200  
cggatactcg ttggtgaact ccgcgaaccg gcgcaccagc cgctcccga ccccgccag 28260  
gtgcgggaac attgccttga tcaaaccgca gtccagacga tcatgatcaa cgagtgtgat -28320  
caccagcagt cactggagcc ggcgttgctt ccacacgggt ggccaaccg agcaggtggg 28380  
tgcacagcaa cgttgttgaa gcacaacgcc tgactggttc aggtccggcg gcagcccgt 28440  
ttccagccgt ctccggcgcg ggacgttctg gttcagagat aaaatcgtgt tgcccgttg 28500  
gtcggtttgg gcggttttga gtccgtgcga ggccgattat gcggcatggg gcctgtgtt 28560  
aggaggtgcc gtggttacta cgcagcgga accggcgggg gctgggggaa cgacggtgag 28620  
tctgccgggc gtaacggccg agttccaccg tcccgatcac tacattcca ctctgatga 28680  
tcttgctagc gcggtgagg ccgttcgggg cgctctccc gaccatcaga cgatggcctt 28740  
ctatcgtggt cttggagcgc tgacctgtt ggggttgatg gactggccgg tggcgcgcg 28800  
gattggcgtg ggtgtcgcg tcgcgcagcg tgccggcgga ccggcgctga agtgatccg 28860  
ttttagcgg gttccgggct aggcgtcgct gagcgcgccg cggcgcgct caggccagg 28920  
gaacctgggt aggtgttcg ttgcctcgct ccggcgtagc tgcatgctcg gagccgtttc 28980  
gaagaacccg agcacacagg cggccttgaa ttcaccaccg ctgggactac tgggtgtgctg 29040  
gcgggtcgaa cagcgtggcg aagcgtcgg tgttgctgat cagctcgag gccaccgaga 29100  
ggacggatgt gcggctgtt cgaggttcgc tttccgctcg tcggtcctac gctcgtgact 29160  
cacggtatct gtgcgcgcgg gaggtagtgg tggctaggac ggcgcggaac acaacgcctg 29220  
gcgcaggcgc gcggcaacaa gggcgcggt tgcgagccgc ggcggtcagc ctgccgctcg 29280  
cggcacagtt tcgtcgggtc gatgttcgca tgccgggtcg ttgggaggac gcggttcgcg 29340  
cggtacgggt cggttgggg ctcgaatggc cggttcggg ggcgaccggc gcgggggtgt 29400  
cacgggcgtc ggtcggcttt gcgtcgga gctgagtgtg ttcgatgctg cgtaccttg 29460  
gcgcgttggc ctgcagatc gtggcgccgg tggtggctgt cgcgcagagc atcgcgggct 29520  
tggggcgcg gcggcgggcg tggctcgtcg agggcggtgc gtacgtcgag gtgcgcggtg 29580  
ttgaccagcc gggggtgag gaggtggcg agcagatcgc ccggcggttg gagaagggtg 29640  
cgggtgttgc gtgggtcgag atcaacgctg tcgtggggcg ggtggctcgc gggcacgac 29700  
cgggcctggt tggggtcgct gagctggtgg gtgtggtgga ggacgcggag gcggacaccg 29760  
gtttggacga ggaaccgttc gctgcggctg gtgtgagcca ccgggacaat ccgggtcaagg 29820  
tgatcgtga ggacaggtg ctgggtgcga acttgaccgg cgctgtgac gcggtggggg 29880  
cccgatgct gccgtgccc gtgtgcgc cgtctgttcc ggcgctggtg tcgaccggtg 29940  
agtcggcctc ctgggtccgg gggccgctgg aaacgcggat cgggaagacc gccactgacc 30000  
tggtgttcgg gtttagcagt gcgttgggga acacgcttgc gcggcgctcg gcgagtttgc 30060  
tggcggaatc ggggtaccgg ttctgcctgt tgcgggagac gcaggcgcg cggcgagct 30120  
gggcgcggtg ggaagccgga acgggggcgc atcctgccag tcaccgtagt gcgcgggtg 30180

ggggttcggcc gcgtcctgtg ccgttgccccg ccggccccggt cgagcgcgcc gcgaacggtt 30240  
 cggctctggg tgggccggcg gcgttcgcgg gcatgctcgg ggtgacgcgg agttttcagc 30300  
 gtgcgcaggc gatggtggtc gcggggcgcg cgcgcgcgcc gcaggtgggg cggggcgcgt 30360  
 tcgcggcgca gctcgggtac gggctgtcgg agcggacgac tctggtgttc gatgcccggg 30420  
 tgttgccggc gttggaccgg gtgcacacgg tggatgatcga cgcggacggt ctgcgcaccg 30480  
 gtgctcgggt ggtgcatgac gtgattccccg cggacggggg gcgcgcggtc gtcgagttgt 30540  
 gggaaactgc gcatgaactc gtgggtgtga ccaagggggt cggcgcgcag cgtgatggtt 30600  
 ggtcgctgta ctcggtcgtc cggtcgcggg tgccccaga cgtgcgcgag gtggccggtc 30660  
 ggggtatcga ggtgttcgcg ctgcgcacag attcgcgcac ggccgcgtgg gttctggctg 30720  
 ccgaggagct ggatccgctc gctgaggagt tggtcgcggc agctcgcgag gtcgggtcgg 30780  
 tggttctcgc gggatgatgg ggccggttgg cggggcggtat cgggtgtgat cgcgtggtcc 30840  
 cgggtggggg gcggctggct tcgggtggtgc gggagatgca gcgggagggc agcgtggtcg 30900  
 gtgtggtgtc cgcgcggggc cgggcggcac tggcggtgc ggatgtcggg cttggtgtga 30960  
 cgcgtggtgg tgaggttccg tggggtgcgc acctgatctc gggcgcgagc ctgggggagt 31020  
 cgtgtgcgtt gttgagtacg gtcccgaccg cgcgttatgt cagcgtgcgt agcgcgcagc 31080  
 tgtcgttgat cggttcggtc agcggggcgg tgttcggcgc gttcgggcct gccatcggga 31140  
 gccaggtccg tgcgggcgtt ccggtggccg tgtgtgcggc aatcgcgctc ggcgtggga 31200  
 cgtggtgggg gatgcaggcc gcgtggcgtc cggttccgcg cgcggtgccg cgcactgcct 31260  
 ggcagtcgat gggccccggc atggtgctgg agctgctcgg cagttcatcg cagggtctga 31320  
 ccgatgatgc ggcgcacgag cgtttccaag gtcaggacga ggggacgggc cttcgccgga 31380  
 tgggtgtgtt gcgggcgtcg ctggaggagc tggcgagccc gttgacgccg gcgttggcgg 31440  
 gtggggccgc gatctcggcg agtatcgggt ccgttgcgga cgcgacgac atcatgggcg 31500  
 tgctggggat gaacgcgttg atcggcgggc tgcagcgcac caccgcggac cgcgctttga 31560  
 acaggcttct ggaggccagc gttgttcggg ttgtggtgcg ccgtaaggaa aatgtgacgc 31620  
 agttgccggc tgacgagctc gtgctcggcg atgtggtgga gttggtggcc ggggatgctg 31680  
 ttccggccga ctgccggctt ctggaggccg acgggctgga ggtcgatgag tcgagcctga 31740  
 ccggggagtc gatgttggtc accaagacgg tgacggcgac tccggcggtg ccggtcgcgg 31800  
 accggacgtg catgatctac cgggggaccg cggtcgctgc cgggcggggc gtgggtgtcg 31860  
 tggtcgcgac cggcgcgaac accgaggcag gcagcgcgat cgggcccgat gcccggcgag 31920  
 gtgccgatgg tgtcgcggct cgcctgcaga cgttgcccaa ggtgacggtg ccgatctgcg 31980  
 tgggagcggg tgggattctg ctcgtcacgg acatgctgcg taggcgtcca ttggggacat 32040  
 cgctggggcg ggcggtgagc ctggcggtgg cggccggttc cgagggattg ccgttcggtg 32100  
 ccaccgtcgc cgaactcgcg gccgcccgcc ggtgtgcgaa acgcggggcg ctggtgcgta 32160  
 actccgcgac gatcgaggct ttggggcggg cgcaggtgtt gtgtttcgac aagaccggca 32220  
 cctcaccga agggcgcac cgcgtgcggc aggtatccga tggcgcgtcg agccggtggc 32280  
 tcgaggaat ctccacagtg gaacggcgag tgctcgcggg tgccttgcg gcaagtccgt 32340  
 tgggtgagaa cggagaactg ctccgcacat ccaccgatcg cgcggtcgtc gatggcgcgc 32400  
 gtggcgcggg agtgaccccg gccgaaggca agcagcgtg ggtaccgatc gacgagatgc 32460  
 ccttcgagcc ttcccggtgg taccacgcgg tgctcggcac ctggcgcgac ggcaatctgc 32520  
 tgagtgtgaa aggtgccccg gagatcgtgc tggcccagtg cacccggtgg cgcagatcag 32580  
 atggcgacgt gccgttggag gagaccgctc gtcgaggcgt cgagcgagtt gtcgaaaacc 32640  
~~tegeccgga gggctaccg gtgctggtg tegeggagcg ggeeggetg gaegeagtg 32700~~  
 atctcgacga gtccccgata cgggatctgc ggtgatcgg tttcgtcgcg ctggccgacc 32760  
 cggtgccggc cactgcggcc gccgccgtcg cgcagctcca gcaggcagg gtcgaggtcg 32820  
 tgatggtcac cggcgaccac ccagcaccg ccgaggcgat cgcgcgcgaa ctcgacgcgc 32880  
 tcaacgagcg tcgcgtcatg accggcaccg aactggacac cctcgacggc gaccaactgg 32940  
 cggcggaagt gaccaccgtc gcggtgttcg cgagggtcag cccggcgcag aaggcgcgca 33000  
 tcgtcgacgc cctgcaacgc aacgaccgtg tggtcgcggg gaccggggac ggcgccaacg 33060

acgccccggc gatccggctg gccgacatcg gcacgcgcgt gggcgagcgc gcaacccccg 33120  
 ccgcccgcga ggccgcccgc ctcgtggtca ccgatgaccg catcgagacc atcgtcgacg 33180  
 ccattgtgga aggcgcgcgc atgtgggcat cgggtgcgtga ctcgttgcc atcctgctcg 33240  
 gcggcaacct cggcgagatc ggcttcaccg tcgtcaccgg tgtgatcaac ggcgggggca 33300  
 gcctcaacgc ccgccagttg ctgctggtca acctgctcac cgacatcctc ccggcgatgg 33360  
 cgatcgcggt gcgaccaccc ccgacgatca ctgcggagat gctgctggca gaaggccccg 33420  
 aagcctcgct gggcgctgcg ctgaccgcgc acatctaccg ccgcgcgcgc gtcaccgcgc 33480  
 gggcgcccat ttgtgctggtg atcctgggca gggcgaccgg aacctgctgc caggccaaca 33540  
 cggtcgcgct cgtggcgctg gtcgccgccc agctcggaca gacctggtc gtgcgcggac 33600  
 ggacaccact ggtcatggcg tgtgcgcttg tctcgttctg ggcaactggcc gctgctgctc 33660  
 aaactccccg cctgagccaa ttcttcggag ccagccccgt actccacac ggctggatga 33720  
 tcgccgctgg cgcaacgacc gccgccaccg tcatcagcct cttgctccaa caatggtgaa 33780  
 cctccatgca tgagcgtcaa gatcccttca ccacagccat cgtcggggccc gactcgggac 33840  
 aacttcgagc aggaagacg ttggagcgcc gagagcgggtg aacgactcgc cagcggcgaa 33900  
 agcgccctgcc ccagccgac gggcggtgga ccacaatgg gagctgcggc gacgtgcatg 33960  
 cggtcggtca tcaggatctc tgtcgtgagc ccgtccggat agatgaccgg cactgcggg 34020  
 taggactcgg cgatcactag gtcgtagtcg tcgacctgga gacctgaag agccgcatcc 34080  
 ggctcggctt gaacaatcac cacatccagt cctggacgcc gccggcgtag acgagtcagc 34140  
 acctcgggga acaggatgaa ggtgacggtc tgcattggagg ccacacgcag acgttgcgag 34200  
 atctcgccac gtgagcgatt gatctgagtg cgcagcgtct ccagctcagc gatcaccgg 34260  
 tccgctgcct ccacgacgag ctccaccttc gccgtcaggc ggacgcgacg gccttcgcgc 34320  
 accagtagcg gagcggatgc cgcaccacgt gccgcgcgag ctggccccgc gcataaccgc 34380  
 cagcagtggc ctgaccagtg ccaacatcgc ccacacgggc ccggtcggcg accttgatca 34440  
 gtggtacagg ctgggcaggc cagcgggcaa cgcacgccct cccgcacccc accacgatca 34500  
 acatcggccg agatcgtaac gaaccccgtc ggtgggttcg aagtgggttcg aacagttggt 34560  
 caagctgtaa gtgcgctgcc agccatgacc gggctacaat aaaacgctag ttctaccgag 34620  
 gtgcggtgag tcgtgtctgg ggttgcggtt cgggaggctg gtcgcctccc atatgaggtg 34680  
 acgagctttg tggggcgccc tcaggcggcc acaagagtta agcgcgcgct gtccagttca 34740  
 cggttagtga cctgactgg ttccggcggg atcgggaagt cccggctggc gctgcacgtc 34800  
 gcacacgacg tgcggcgggc cttctccgcc gggacatttc tggtgagct ggctaacgtt 34860  
 cgcgatcca tgctggtgcc gcaggcggtc gcggctgctc ttgggatcca tgaccgttcg 34920  
 acacgtgccc cggaggccgt gtcgtcgat tacttggcgg acaagcaagt cttgctggtg 34980  
 ttggacaact gcgagcacct gctgggggcc tgcagccgcc tagtcgctgt gctgctggcc 35040  
 gcggccccgc ggttgccgat acttgcaacc agccgagaac atttagggat agtggcgga 35100  
 cagacctggc cggttccgcc actgtccgtg cctgctggca gcgattcacc aacgtcaca 35160  
 caggaaggtc gtccctacca gcgtgaagcg ttgacctgt tcgagcagcg tgccgcggcg 35220  
 gcgctaccag ggttcacat caacgaagaa aacgaacaga tcgtggccag gttatgccg 35280  
 aggttgacg gcattccct agcgatcgaa ctgcgcgcgc tgcgcgttcg cgtgctctca 35340  
 gtcgggcaaa ttctaaccgc aatggagaac cgattccggc tgctcaccac cggcgaccgc 35400  
 gccgacgat ctcggcatca gacctgcgc gccgcgcgtc actggagcta caacctgtgc 35460  
 aatgagcatg agaagcagtt gtgggcgcgc tgttcggtt tcaccgacga attcgatctc 35520  
~~gaegcageeg aaaacgtctg caecggcgac gggctgaecg atgatgaegt ettegeegct 35580~~  
 ctcgcgggac tgatcgacaa gtccgtgctc acccgagccg aggagcgcac gaagacgaga 35640  
 taccggatgc tggagactat ccgtcaatac gggcacgacc ggctcaccga ggcggtgccc 35700  
 gaaaacgctc tgcgctgccg gcaccgcgac tactacctgc atctcgccga gcagtccgat 35760  
 gccgagtcgc cgggcccgc tcaagccggt tggctcgcgc gtctgcgcgc cgagcgcggc 35820  
 aacttctggg cagcgttgga ctactgctc accacacctg gcgagcgctc ggttggttg 35880  
 cggatggtga gcgcgctgtg gttctactgg gctgctggcg actatctgag ggaggccgc 35940

atgtggctgg acagggcggtt ggccctggag accgaaccca ccagccagcg tgcccgcgca 36000  
ctgtggatca ccggatggat tgcccaccgg caaggggacc gagatggggc gctggtcctg 36060  
cttaacgaaa gccgcgagct ggcccggaaa ctgggcgacg aaacggaatt tacctacgcg 36120  
actcagtttc ttggcgaccg ggagatgtgg gacaacaacc tgactcgcg caccgcaactg 36180  
ctcgacgagg cactggcccc ccaccgcgca actggccact ggaccggccc cgccttagcc 36240  
atcttctcaa tacgagcgca ggcagcgggc ctgcttggag acctgaccg ggcgatggtc 36300  
cttctgcacg aatgccgcaa gatctgtaca gaacttggcg agcgtggag gctatcgtag 36360  
gcggagtggg acatggcgat cacttggtag gccgcggcg acccaaccaa ggcagccgcc 36420  
tccgcgtcca ggtccctccg caataagcgg gaactcaacg acctactcg gatcagcgac 36480  
tgctgtgaac tgctggcctg ggtcgccgcc gcagaacgca atcctgaacg cgcggcgatc 36540  
cttttcgggg cgttggataa aatgtgggaa ctgattggca ccccgctgtt cggatccgag 36600  
acgctactga cctggcgtga gcaggccaag gcacgcgccc aggaagcact cacagacacc 36660  
gaatacgaia ccgcccagag ccaggggacg cggatgaacc gggagcagac tatcttctat 36720  
gctcttgggg agaagccacc gccggtcgaa gcggccgccc cttcggctcc ggaaaccgaa 36780  
ccggttctca caaacgaga acgcgaggtc gctgccctgg tcaccgctgg caaaacaaac 36840  
aaagagattg cggctgacct ggtcatttcc cagcgcacag ccgaggccca cgtcgagaac 36900  
atcctttcca agctgggctt tacctcgcg gcacaggtca tcagttggat gaccgggagc 36960  
agataacgta gtcacctgct ctgttggccc atgcccttct cgggcccagc cggctcacat 37020  
gatccgttcg cggcccggca acaccgccc tctgtggcaag gtctgcgata ccgaaagcgg 37080  
acatccgctt cagctaatac tccgcggccc ccgcaacggc gcatgttgaa gtaggtatgg 37140  
tccatgatcg gaggttggcc attgccgatg gccatgaccg cggagttcgg gtgctgcatg 37200  
tagttgcggc cgcccacata ggacccttg gcgagcttac gcggtataga ttgagccacg 37260  
ggccagcttc gtaggccttg gctgttgccg ccgacgagta gtgagtagcc ggtgtgaacg 37320  
gtctcccagc accatcgagc caccgttcgt ccggttcta gtgcccgctg gcgaagcctt 37380  
gcacacgtcg atagcgggct cttgagccaa aaactcgtcc cgttccaaca cgagcacatc 37440  
gagtcgcgca ccagcgagcg ccccggcgac gatcagtcag cctccaccgg acccgacgat 37500  
cgccacgtct gccgatgccc ggttgtcctt gtgcatgac tggtgtgcct aactgcgccc 37560  
agctttcgtc atcgagcctg tcccacagtc tgccgatgcc cggtcggagc agcctccagc 37620  
gccgcccgtg cgaaggccaa cgaccagacc ccgtcttggc cggaggccgg gggctcggcc 37680  
tcgccacgca cagcctcgat gaaccgtcga acaccacga cgtagaggte ctcgcggtta 37740  
cccagatcga tcttcttgga cccggccgga ccgacgtgtg tcagcgtgcc gaccggtgac 37800  
atcgactgga taccgtgcc gatgagcgt ccttcggttc cgaaaacctc caggcatgtg 37860  
ggggcatgtc cggtcacgaa gctttccacg aatgaagcat gcacaccgct gacgaactcc 37920  
gcgcttccgg agattgtct atcgattccc accctcgtca tgtggccgga cgaggtatgg 37980  
gccacgaccg ttagcggatc ggcattgagt acgaaccgga gacagtccgc gtcgtgcacc 38040  
gtgaggtcga gtgcgactcc ggcgcccgt gccgaatcgg tcgtccgcca tctccgcg 38100  
gcgggcccga gtcggatcgc gttccggacg ttcactgca ccggctcgcc gatgtcgccg 38160  
gcggcgattg catcccgat caaacgatgt ggcacggagt tgcgcatgtg gtggtttgtc 38220  
gcgaacacta cgctgcctt ctccggcgca tcgacaagtt cttgtgctc ggcgacggtc 38280  
agggcaagcg gtttttcgca caggacgtgc ttgcccgtg cgatcgccgc aagtgttgga 38340  
tcgtgatgtg ccgagttgat actcgagatg tacacggcgt cgacgtcagg atcgctaagc 38400  
gccacggtga gttcgtcatg gcccggggc ccgaatggtg ccgctacatc ctgctcgtg 38460  
gccaggtcgc cgctgacgac gctgaccacc gcggcgtag gctgggcccga gatcgctggg 38520  
cctacaaatg cgcggcgat gtcgctggcg cccacgagcg accagcgga gggacggttc 38580  
gtcatgattc gatgggttcc tttcgggaat accgcgctg tgcgacattt ccgctggcg 38640  
acagacgtga ggccagcgcg ggtggccgag aggattacac gaccatgaaa cgggtcgccg 38700  
cttgactgtt agcaagatcg gaggtttcgc ataccgccg tgatctcgt tggccacgag 38760  
ctgtcctggt agtgccaggt cgtcaggatt ttgtcagcgc tcttcccca cccggctgca 38820

cgtctcccga gccggccctc cggatcgaac agccccgacat cgccgtcggc tggctggatg 38880  
 ccgtagcgctc ggcgtagctg tttccaggga gtacgccggt gtttgcgctc ctgccaccgg 38940  
 acgacctctt tccagagata cgatcgtagg tagctgaagg tcgcgttcgg gcatccatac 39000  
 ctgaagtacg ccttccaggc agcatccggt tccggatagc tttcaatgtg gggaaaagct 39060  
 gggctgcggg cgaccgaaa accgacggca ccgctcgcg gcagcgtgca ccctggaggc 39120  
 gaggtccgtg ctgctgatcg tgatcccgt ggcgagtgg tagccctgga tgaccgtaga 39180  
 tcagcgcacg taggtgatga gtttgggctc gggatcgctg acgtcagtga ccccgaggag 39240  
 cggacatggc ttcgggttgc gctgatctgg agacaggatg tttgtcgggg ttacacgttc 39300  
 tggagcagtc cgccatcgac acggacaacg gatccggtga ttagctggc acgtcgctg 39360  
 gcgagaaatg tgaccgcgtc ggcaattcc tcgcgccggc cgtagcggcc gaccggaata 39420  
 gagctctctg attccctcgc gatgttctcc gcagtgcggt tctcccgctc ggcctttgcc 39480  
 tggctcaact gggcgatgcg cgcagtggcg atgcgcccc gaacgacgat gttgctggtg 39540  
 atgcgctcgc ttgccacctc accggacagt gtcttgacc agcccaccag actcgagcgg 39600  
 agggcgcttcg acaacccgag gttggggatg ggcgcatga cgccggacga agtcgacgtg 39660  
 atgatccggc cccatccccg ctcggtcatg cccgggagca cgcggtctgt gatcgcgacg 39720  
 acggagagga tcatcgctc gaagtgccga cgccacaggg tggcgctctg cccggacaca 39780  
 cgggtggcg gcgggcccgc ggtgatgttg accaggatct ccaccgggccc gagctcgctc 39840  
 tcgacggcgg tgacgttctc atcgatgacc gacaggttcg ccaagtccca gaccaggctc 39900  
 atgttatcgc cggatcatga cgcctcgacc tccgtgagcg cgtccttctc gatgtcggcc 39960  
 agcgccacac gcgcgccttc ccggcttagc gaggcgga tggcgaacc gagtccgccg 40020  
 ccggctccga gtacgagggc gactcttcca ttgattgcga gatccatgac atgtccttgc 40080  
 ggtttcgggt gacggcggcc cgaaagtgc tggcacgcc tgatggagg gacattggtg 40140  
 ggaggagtgg cagtggccc gtcagccgc taccacggcc tgagcgtga cctcgaacag 40200  
 gacaccgggt cggagaccgg cggatggtg cgtccgggccc gggccgggct cgctgaagta 40260  
 ctgcggttag atcggttca tcgctcgaa gtcgtccccg tcagcgagga acacattcac 40320  
 ggtgatcacg tctgcctcgc tgcagccggc tgaacgcagc gcggtcatga ggttctcgaa 40380  
 tgcccggcgc acctgaggtt ccagaccacc cgggtcgagg gatctgtcgg cgaggtagcc 40440  
 gcactgcccg gcgattgcca cgatgttgcc gatgcgtacg gcggcgaggt aagggcctcc 40500  
 cgcgggagga aggtcgacgt ctgtgaacgc ctgacgggac acggtgtcgg ggcctctctg 40560  
 tggctggtgg tcatgtccaa ctgacttcgg gagcttgtgg ggtcggttca accggagcct 40620  
 cccgtaggag gtcttgagcc ccgtcagaag aacgtgagga cgccgtcgac gacgttgaag 40680  
 tcgtcgtcga tgatcgggat gaaggggagc ttgtcgaagg ccgtgcacgg gtgggacatg 40740  
 tcgaaagcga cgatgtctcc gggagcgatc gcgagcccgg gggcaacgcg aacgaaggcg 40800  
 tgctggtcca tcagcttgaa cgtctccgct tcccgcagcc gctgcctcga gccatcccg 40860  
 cgaaacagct cgcgtggctg tggcagcgtc aggtcatggg ccgcgtcgcg cttcccgatc 40920  
 gacaggatcg cgagatcagg ctccggccgg gacagcacgg aagcccatgc agtgagtgcg 40980  
 ttgtggagcc ggagcggttc ctgctcggcc cgtcggccag cgagcggcga gagctttctg 41040  
 taattgcccg cgtcgtggga gatgtaacag ccgttccgaa gcactaggtc cagctcgtcg 41100  
 cggacaccgt cccacgagcc gagtccggcg acgacacgat cgaagtacga gcttccgctt 41160  
 gccgtcacga tgatgcggtt ggcgtcgaac aggtcgcgac gtgccagctc gacacgatac 41220  
 gaccggacgc cagcgaagaa ggtgtcagat gcatggaggt cctccggtgt catgccgccg 41280  
~~gtcacgagc cctcgtaggc ctgcagccg accagcggga gatgctcgtt cgcggcgacc 41340~~  
 gccgcggcga cctcgagcgc cgagtcgagg tctcgacgc cacacctgcc accaggcact 41400  
 ccgacctcga ggaggacgtt gagggtgacc ccgtggccgg agtcggcgat ggtatcggac 41460  
 atgcgccgca ccgtgtccgc atcgtcgacc aggcagtaga agtcgaacga gctgtcgtcg 41520  
 gcgatttcgg cagtcaccca acgcagcgc ggagcatcga cgagttcgtt ggccaggatg 41580  
 atgcgctgaa ccccgtagcg gcgcatcacg gccacctgag tgggcgtcgc tgccgtgaac 41640  
 gccacgcac cgtgagacgc ctgaagggcg aagagctggg gtgacatggt ggtcttgccg 41700

tgcggggcga	gaagcgcggt	gttgcgctcg	cagtacgcct	ccatagtcgc	caggttggcg	41760
ttcatagcgc	gtgcgtggag	tgtggtcacc	ggcagggcga	ggtctccgcg	ggccacgttc	41820
cagttctgct	ggteccacct	tccgaggagc	acacgagcac	ccaggggcac	gcccttgagt	41880
gcgtcgtcga	ggacgggtct	gcggagggtc	tcggcacatt	gcaggagtcg	ggtccgggtc	41940
agctgctcga	ccgcgtgagt	gctcatcgtg	tagtggcctt	tcgcgccgtc	gtggtcagct	42000
gaccgatgcc	gtcgatccgg	cactccacac	ggtctccgtc	gccgatgacg	agcgcgccag	42060
gggttccggt	ggagatgatg	tctcccggaa	acagcggcat	catcttcgag	tggaagctca	42120
cgaggctggc	gggcttgtgg	gtcatgtggg	cgaccgtgtt	ggtcctgaac	tcatcgccgt	42180
tcttgaccgt	cgagatctcg	atgtcgtcca	agctttcga	cttctctatc	acctcatcga	42240
gcggtacgag	ctccgggtcc	aagcagaaga	acgtccggaa	gttcttcgag	cgggtgagat	42300
agcgcggtt	gatccgcaag	atgtcctcgg	cggctctggtc	caggatggtc	gtgaccccc	42360
acacgtagtc	catcgccctg	gactcgctga	cgttctcgca	agtcttgccg	atgatgagcc	42420
cgatctcagc	ctcggacgtg	gtgcgctcac	tctgcgccgg	aatcacgata	ggctcatccg	42480
ggccgatgat	cgtgtggtcg	cacttcacga	aggacgcggg	ctggtcgggc	gccgactcgc	42540
tcaggctccg	tgcgtgatcg	ccgtagtcca	gaccgatgcc	ccagatcttc	cggggggatc	42600
ggtacggcgc	ggtgaaccgc	acctcgtcca	cgcggtgtaa	ggcgagggtc	ggggcggttct	42660
tgacacgctc	acgcagccgc	gcgagggtgt	ctcctgcgat	cacgtccatg	atggctcgtga	42720
ccgccagctc	ggggctctatc	tcccggattg	ctgcgacgcc	gcggtcgctc	acgacaacgg	42780
ggatctctcc	ttccgcagcg	gtgatggtcg	ccagggtgctc	ggtcacgaaa	gcccgctcct	42840
cggtcgaatg	gtgaggtagg	gggtggggga	gttggtttaa	ctcgttcccg	ggggcacgtt	42900
cccgcgctgc	gtgtgtgagg	agaggacaag	cgacatcacg	tggaggatgt	ggtcggcggt	42960
gttccgggct	gcccggtcgg	cgtcgccagc	cagaacggcg	gatgccgtcg	cgatgtgctc	43020
cgaggcgag	gagatggagt	caccgggtcga	gggagtgtcg	cgttctcgga	gtacagcggt	43080
cagcaccagg	gcgttgctga	ccgcgcgctg	gaggatctcg	ttcccgagca	actcggcgat	43140
gatgtggtgg	aagacgcgat	cctcggcgcg	gaactgggcg	gagtcgccgg	actctgccgc	43200
gcgcagggat	gagggtggcg	cctcgtggag	cctctcccg	tgggcgctgg	tagcccgatg	43260
tgcggtgaac	cagggtgccg	agctctccag	ccccgcccgg	tgctcgtagg	cgtcgcgaat	43320
ggcgcgacc	gaggcgctga	tgacgcgcag	ccccttcggg	gcagcctcga	ccaggccac	43380
gtgccgcaag	cggagaagcg	cctcgcgcac	gggggtcttg	ctgacattca	gctcggccgc	43440
gagcttgcc	tcgctcacac	gacgccag	tgggagtgc	tgacttacga	tgctgtcgcg	43500
gatggcctcg	aagaccgat	ctgtcagatt	gtcgacgcgt	gcgtcgagg	gtgcgagcat	43560
accgttgccg	gacccaccg	cggagtctgc	ggcggtggg	cgatggtgca	gctccaccat	43620
gttcagcgtg	cccgcagtgc	gtggctgatt	cggttccatc	gtcccccttc	cttgacagt	43680
ggaatacgg	atatggtcct	caccgcgtaa	gcaaccgcc	gaggcgggac	tccgtagcgg	43740
atgaagtcct	tgggatatgt	cgtataccga	gggagagctg	atgaagccc	agtcgcaccc	43800
ggccttcgag	agcctggatc	gactggccac	gatcgagatg	cgtccgccga	cgagcccggc	43860
cggcatcgtg	agcccgcttc	acacgctggc	gcggggtgag	cggtcggagc	cgatcagcct	43920
tgcgatcgcg	cgctctctgc	tgcgtcgaca	gggtgaaaag	gcgttgatcg	tcacgggtct	43980
ggtcgaggaa	ggccggttcc	cgcaggggga	ggtggacggc	cccatcggt	ctctggcact	44040
cgcgcgcacg	ctcgcgggtc	ttggtagcga	ggtgaccatc	gtcatcgacc	cagaggccgt	44100
cggcccggtg	gagaagctgg	tcgacgctgc	gggcctggat	ggcgtgacct	tgatggagtc	44160
<del>cgatctctcc</del>	<del>agcgcgccg</del>	<del>aggeecggga</del>	<del>gtteggggg</del>	<del>eagttegggg</del>	<del>tegteggcgc</del>	<del>44220</del>
catcgagaag	ctggggcaga	actctgttgg	cggccgtcat	ctgatctggg	gcacgccggt	44280
caccgccggt	gatctgtttg	ccgacgacta	ccttcgcgcc	gccacggaca	acggtgctct	44340
gacgctggcc	gtgggagaca	acggcaacga	gatcgggttc	ggcaacattg	ccgccgagtc	44400
cgaggcgctg	actccgcgag	gcgtcagcgt	tgaggggcga	ttcttcgcgt	ccacgacggt	44460
ggatcatctg	ctgcctgcat	ccgtgtcgaa	tttgggctgc	taogtcatca	ccggtgcggt	44520
ggccatcctg	gctcgacgcg	caaacctggc	agtcaccggg	gatctggtgc	gcgagtggac	44580

cgagctgggg ctgcgcggt gtctacgcag cggcggcggt gacgatccg cgttccaagg 44640  
 cgacgacggg atccactgc gtttcgtcgc cgcgcattgc gagctgatct ccggaatcgt 44700  
 tcaccagagc ctgctcggcg acccctgggt ggcagaccag tgacggccac ccgtagcggg 44760  
 caaccctcgg ttgccagcct tcgtcatcat cgcgcgggaag ggggaagtcac cgatggagtt 44820  
 cgttgatctg cagcacgcgg agcccatcga ggcattggcgt gccaaaccgc cgaatgcgtg 44880  
 ggcgaggccc acgtccggga tctgtcatgg attcgtgcag gccaaacctg tgatgctgcc 44940  
 ccaggatgac gcgttcgact tcatgcgggt cgcgcagcgg aaccccaagc cgtgcccgat 45000  
 cctcgaggtc accgacgcgg gagatccgga gccgaagctg accgcccctg gtgctgacct 45060  
 ccgcaccgac ctgcccacaa acagcgtgta ccgggaaggg gcgctgggtg aggagcgcag 45120  
 tgacgtcgtt gacctgtgga ccgacgacat ggtcgcgttc ctgcttggtc gcagcttctc 45180  
 cttcgagcac ttgctgctcg atgccggact cccgatccgg cacatcgacc aggggatcaa 45240  
 cggcccgggt tacctcagca accgtgagtg tgtgcccgca ggccagttcc gggggaagct 45300  
 agtggtttca atgcgtccca tgcccccgga ccgggtggcg gaggtgtca ccatcacgca 45360  
 tgcgcattcc gaggtgcacg ggggccccgt gcacatcggc tcccctgaga agttgggcat 45420  
 cgacaacgtt gatgaacctg actgggggtga acgagtgcgt atggagccgg gcgacgtgcc 45480  
 ggtgttcttg gcctgcggcg tgaccccgca gcaggcggcg atcgaggcga agaccgagct 45540  
 aatgatcacg cactcgacgg gtcacatgtt cgtgaccgga atgcggatcg actccgagtg-45600  
 aggcgggggc cccgggggtg tgcc 45624

<210> 5

<211> 50000

<212> DNA

<213> *Saccharopolyspora spinosa*

<400> 5

gggcgcgccg caattcgatg acgttcatgc gccgtgtcgg ggaatcgccg gtggcggcgc 60  
 cagcagaggc tgaacttact ggtggtgtgt ccaggaaatcg gaggggcagt accgaatgag 120  
 cgaagccggg aacctgatag ccgtcatcgg actgtcctgc cgcctacccc aggcgcctga 180  
 cccggcttcc ttctggcggt tgctgcgcac cggaacggac gccatcacca cgttcccgga 240  
 agggcggttg ggcgaccctg tgctgggtcg ggatgcgccc aaggggcccg aatgggggtg 300  
 tttcctgggt gatgtcgact gcttcgatcc cgagttcttc gggatctcgc cgcgagaagc 360  
 ggcagccgtg gatccccagc agaggctggc tctggagctc gcctgggagg cactcgaaga 420  
 gcgcgggtat cccgcggcg agctgcgcgg tactgccgcc ggtgtgttca tgggggcgat 480  
 ctctgacgac tacgcgcgcc tgctgcgcga gagccgcgcg gaagtggctg cgcagtaccg 540  
 cctcaccggc acccatcgaa gcctgatcgc caaccgcgtg tccatgtgc tggcctgcg 600  
 cgggccaagc ctgacgggtg attcaggtca gtcctcgtcc ctggtcggcg tgcattctgc 660  
 cagcgagagc ctgcgacggg gtgagtgcac gatcgactc gccggcggcg tgaacctcaa 720  
 cctggccgcc gagagcaaca gcgctctgat ggacttcggc gcgctctccc cggacggctg 780  
 ctgcttcacc ttcgatgtgc gggcgaacgg ttacgtccgt ggtgaggcg gcggccttgt 840  
 cgtgctgaag aaggccgatc aggcgcacgc cgatggcgac cggatctact gcctcatccg 900  
~~cggeagegeg gteaaacg atggggcggt tgccgggtc accgttccgg cggcggacgc 960~~  
 ccaggcggag ctgctgcgcc aggcataacc gaacgcgggc gtcgaccgcg ccgcgtgca 1020  
 gtatgtogag ctccacggca gcgcgaccag ggtcggggat ccgctcgaag cagcagccct 1080  
 cggagctgtc ctggggcgcg cgagacggcc cggcgacgag ctgcgtgtgg ggtcggcgaa 1140  
 gaccaacgtc ggccatcttg aagcagcggc gggcgtcacc gggttgctga agaccgcact 1200  
 cagcatctgg caccgcgaac tgccgcgag tcttcatttc accgccccca acccggaat 1260  
 cccgctggac gaattgaacc tacgcgtcca gcgtgatctg cggccgtggc cggagagcga 1320



ggggcccgtg ctggccggcg tcagcgcctt cggaatggga ggcacgaact gccacctggt 1380  
 gctctccggc acgtcccggg tggagcgacg gcgcagtggg cccgctgagg cgaccatgcc 1440  
 gtgggtcttg tccggccagaa caccggtcgc attgctgtcg caggcggcgc gcttgcacac 1500  
 gcacctcaat acggccggtc aaagtccgtt ggacgtcgcc tactcactgg cgaccactcg 1560  
 atccgcgctg ccgcaccggg ccgcgctggt cgcggacgac gaaccgaaac tgctcgccgg 1620  
 gttgaaggcc ctgctgacg gcgacgacgc gccacgctg tgccacggcg cgacttccgg 1680  
 cgagcgggca gcggtcttcg tctttcccg acagggcagc cagtggatcg ggatgggtag 1740  
 gcagctgctc gaaacctccg aggttttcgc ggcgtcgatg tcggactcg ccgacgcatt 1800  
 ggcgcgcac ctggattggt ccctgctgga tgtgctgcgc aacgcggccg gcgctgcgca 1860  
 ccttgaccac gacgatgtcg tccagccgc gctgttcgc atcatggtct cgctcgcgga 1920  
 gctctggcgt tcgtggggcg tgcgtccgtt ggcggtcgtc gggcactcgc agggggagat 1980  
 cgcgggggcc tgcgtcgccg gggccctgtc cgtccgcgat gccgccaggg tggtagcggt 2040  
 gcgcagcagg cttctgacgg cgctggccgg cagtggcgcg atggcctcgt tgcagcatcc 2100  
 cgccgaagag gtgcggcaaa tcctgttgcc ctggcgcgat cggatcggcg tggcgggggt 2160  
 gaacggaccg tcgtcgaccc tgggtgcagg ggaccgggag gcgatggcg aactgctggc 2220  
 cgagtgcgca gaccgagagc tccgatgcg ccgattccc gttgaatacg cctccattc 2280  
 gcctcacatc gaggttgctc gggatgagct gctggggctg ttggcgccgg tcgaaccag-2340  
 gacgggaagc atcccgatct attcgacgac gaccggggac ctgctggacc ggccgatgga 2400  
 cgccgactac tggtagcgca accttcgtca accggtgctg ttcgaagcgg ccgtcgaggc 2460  
 cctgttgaa ggggggtacg acgcattcat cgagatcagc ccacaccgg tgctgactgc 2520  
 gaacatccag gaaaccggcg tgcgagcagg gcgggaggta gtggcgctcg ggacactccg 2580  
 ccgcggcgaa ggtggcatgc ggcaggcgct gacgtcgctg gccagagcac acgtacacgg 2640  
 agtggccgcg gactggcacg cggctcttcgc cggtagccgg gcgcagcggg tcgacctgcc 2700  
 gacgtacgcc tttcagcgac agcgctactg gctggacgag aagcttcccg acgtcgccat 2760  
 gcccgagagc gacgtgtcga cggcgttgcg ggaaaagctg cggctcttcgc cgagggcgga 2820  
 cgtggactcg acgacctca cgatgatccg ggcacaggca gccgtggctc tcggccactc 2880  
 cgatccgaaa gaggtggacc cggatcggac gttcaaggac ctgggcttcg attcctcgat 2940  
 ggtggctcag ctgtgcgacc gcctaaacgc cgccacaggt ctgcgactcg caccgagcgt 3000  
 cgttttcgac tgtcctacgc cggacaagct cgcccgccag gtacggacgt tgttgttggg 3060  
 cgagccggct cccatgacgt cacaccggcc ggactccgat gcggacgagc ctatcgccgt 3120  
 gatcgggatg ggctgtcggg ttccgggtgg ggtgtcctcg cccgaggagt tgtggcagtt 3180  
 ggtcgccgct gggcgggacg tcgtgtccga gttcccggt gaccgaggtt gggacctgga 3240  
 gcgtgcgggg acatcgcacg tgcgcgccgg cgggttcttg catggcgccc cggattttga 3300  
 ccccggggtc ttccggattt cgccgcgcga ggcgttggcg atggatccac agcagcgggt 3360  
 gctgctggaa atcgccctgg aagcagtcga acgaggcggg atcaaccgc agcacctgca 3420  
 cggaagtcaa accgggggtc tcgtcgccgc gacctccctg gactacgggc cagcctgca 3480  
 cgaagcgtcc gaggaggcg ccgggtacgt gctcacccgc agcaccacga gtgtggcgtc 3540  
 gggtcggggt gcgtattcgt tcgggttcga gggccctgcg gtgacgggtg atacggcggtg 3600  
 ttcgtcgtcg ttggtggccc tgcatttgcc gtgtcagtcg ttgcgttcgg gtgagtgtga 3660  
 tctggcggtg gccggtggtg tgaccgtgat ggccacgccg gggatgttcg tggagttttc 3720  
 gcggcagcgt ggtttggcg cggtggggc gtgcaagtcg ttcgcggagg ccgcgcagcg 3780  
 caccggctgg teegagggtg etggcctggt tetactggag cggttgtcgg atgcccgccg 3840  
 gaatgggcat gaggtgctgg cggttgttcg ttgtagtgcg gtgaatcagg acggtgcgtc 3900  
 gaatggtttg accgcgccga atggttcgtc gcagcagcgg gtgattgccc aggcattggc 3960  
 gagtgcgggg ttgtcgggtg ccgatgtgga tgcgtgtggg gcgcattgga cgggcacgcg 4020  
 gcttggtgat ccgatcgagg cgcaggcgct gatcgccacc tacggccagg gccggcttcc 4080  
 ggaacggcca ttgtggttg gctcgatgaa gtcgaacatc ggtcacgcgc aggcagctgc 4140  
 ggggatagcc ggcgtcatga agatggtgat ggcgatgcgg cacgggcagc taccgcgcac 4200

gttgcacgtg gatgagcga cttctggggt ggattggtcg gcggggagg ttcaactcct 4260  
 tacggagaac acgccctggc ccgggagtggt tegtgttcgt cgggtggggg tgtcgtcgtt 4320  
 cgggatcagt ggtactaacg cgcacgtcat cctogaacag cccccgggag tgccgagtca 4380  
 gtctgcgggg ccgggttcgg gctctgtcgt ggatgttcgg gtggtgccgt ggatggtgtc 4440  
 gggcaaaaca cccgaagcgc tatccgcgca ggcaacggcg ttgatgacct atctggacga 4500  
 gcgacctgat gtctcctcgc tggatgttg gtactcgtg gcgttgacac ggtcggcgct 4560  
 ggatgagcga gcgggtggtg tggggtcggg ccgtgaaacg ttgttgtgcg gtgtgaaagc 4620  
 gctgtctgcc ggtcatgagg cttctggggt ggtgaccgga tctgtggggg ctggggggccg 4680  
 catcgggttt gtgttttccg gtcagggtgg tcagtggctg gggatgggccc gggggcttta 4740  
 ccgggctttt ccgggtgttcg ctgctgcctt tgacgaagct tgtgccgagc tggatgcgca 4800  
 tctggggccag gaaatcgggg ttcgggaggt ggtgtccggt tcggatgcgc agttgctgga 4860  
 tcggacgttg tgggcgcagt ccgggtttgtt cgcgttgacg gtgggcttgc tgaagttgct 4920  
 ggattcgttg ggggttcggc cgagtgtggt gttggggcat tcggtgggcg agttggcggc 4980  
 ggcgttcgcg gcgggtgttg tgtcgttgc ggggtcggct ccggttggtg cgggtcgtgc 5040  
 ccgggtgatg caggcgttgc cgtctggcgg tgggatgctg gcggtgcctg ctggtgagga 5100  
 gctgttgttg tcgttgttg ccgatcaggg tgatcgtgtg gggatcgccg ccgtcaacgc 5160  
 tgcggggtcg gtggtgctct ctggtgatcg ggatgtgctc gatgacctt ccggtcggct 5220  
 ggacgggcaa gggatccggt cgaggtggtt gcgggtgtcg catgcgtttc attcgtatcg 5280  
 gatggatccg atgctggcgg agttcgccga attggcacga accgtggatt accggcgttg 5340  
 tgaagtgccg atcgtgtcga ccttgaccgg agacctcgat gacgctggca ggatgagcgg 5400  
 gcccgactac tgggtgcgtc aggtgcgaga gccggtccgc ttcgccgacg gtgtccaggc 5460  
 gctggtcgag cacgatgttg ccaccgttgt cgagctcgtt ccggacgggg cgttgtcggc 5520  
 gctgatccag gaatgtgtcg ccgcatccga tcacgccggg ccgctgagcg cggccccgc 5580  
 gatgcgcagg aaccaggacg aggcgcagaa ggtgatgacg gccctggcac acgtccacgt 5640  
 acgtggtggt gcggtggact ggcggtcgtt cttcgccggt acaagggcga agcaaatcga 5700  
 gctgccacc tacgccttc aacgacagcg gtactggctg aacgcgctgc gtgaatcttc 5760  
 cgccggcgac atgggcaggc gtgtcgaagc gaagttcttg ggcgcgctcg agcacgaaga 5820  
 tgtggaatcg cttgcacgcg tattgggcat tgtggacgac ggcgctgctg tggattccct 5880  
 gagaagcgcc cttccggtgt tggccggttg gcagcgaacc cgcaccaccg agtccattat 5940  
 ggatcagcgg tgttaccgaa ttggctggcg gcaggtagcc ggactccgc cgatgggaac 6000  
 tgttttcggt acctggttg tcttcgcgcc tcatggttg tccagcgaac ccgaggtggt 6060  
 ggactgcgtt acggcactgc gggcacgttg tgcctcgtg gtgttggttg aagctgatcc 6120  
 cgacccgacc tccttcggcg accgggtacg aaccctgtt tcgggccttc cgatcttgt 6180  
 tggcgtgttg tcaatgttg gcttgaaga atcggtcctt ccgggatttt ctgcggtgtc 6240  
 acggggtttt gcgttgaccg tggagttggt gcgggttttg cgggcagctg gtgcgactgc 6300  
 ccggttgttg ttgctgacgt gtggtggcgt gtccgtggga gatgtaccg ttcgtccagc 6360  
 gcaggccctg gcgtgggggt tggggcggtg tgtggggttg gagcatccgg actggtgggg 6420  
 cggcttgatc gatattccgg tcttgttcga cgaagacgt caagagcgggt tgtcgattgt 6480  
 gctggcaggt ctgatgagg acgaggtcgc gatccgtcct gacggcatgt tcgcgcgtcg 6540  
 gttggtacgc cacactgtct cagctgatgt gaagaaggcg tggcgcccca ggggatcgg 6600  
 gctggtgacg ggcggcacgg gtggtttggg ggcgcacgtt gctcgtggc tggccgacgc 6660  
~~cggagccgaa catgtggga tggtagtcg acggggcgag caggacgga gtgtgagaa 6720~~  
 gttgcggacg gaactggagg atctgggtac ccgggtgtcg atcgtgtcat gcgatgtgac 6780  
 cgatcgcgag gcgtcgcgg aagtgtgaa agcccttcgg gctgaaaacc cgttgaccgc 6840  
 ggtagtgcac gcggcaggcg tgatcgagac tggtagtcg gcggcaatga gcctggctga 6900  
 tttcgatcac gtgttgcgg caaagggtgg ccggtccgcg aatctggatg cttgttggc 6960  
 cgatgtggaa ttggacgcgt tcgtcttgtt ctcatcggtg tcaggagttt gggcgctgg 7020  
 gggacacggg gcttacgcag cggcgaatgc ctatctggat gcgctcgcgg aacagcgtcg 7080

gtcgcgaggg	ctgggtcgca	ctgcggtggc	ctgggggccc	tgggcccggc	agggcatggc	7140
ctccggagaa	acaggagacc	agctgcgccc	atacggcctt	tccccaatgg	ctccgcagca	7200
cgccatcgcc	ggaatccggc	aggccgtgga	acaggacgaa	atttccctgg	tagtgggcca	7260
tgtcgattgg	gcacgtttca	gcgcgggatt	gctggcggct	aggccgcggc	cgctgctgaa	7320
cgaactggcc	gaggtcaagg	aactcctcgt	cgatgcccag	cccagggcgg	gagtccttgc	7380
cgacgcgtcg	ttggaatggc	ggcagcgatt	gtccgcggca	ccgaggccga	cacaggaaca	7440
gctgatcctg	gagctggtac	gcggcgaaac	cgctctggtg	ctgggacacc	ccggggcagc	7500
ggccgttgca	tcggaacgag	ccttcaagga	cagcggattc	gactcgcagg	ccgcggtcga	7560
actccgcgtt	cggtcgaatc	gagctaccgg	cctccagttg	ccatcgacaa	ttatcttcag	7620
ccatcccacg	cctgcggaac	tggctgcgga	gctgcggggc	aggcttcttc	ccgagtccgc	7680
aggagcaggc	attcccagag	aggacgaggc	gcgaatcaga	gcggcactga	cgtcgatccc	7740
gttcccggcc	ttgcgcgagg	caggcttggg	gagtcgcgtg	ctcgcacttg	ccggacaccc	7800
ggtcgactcc	ggtatctcct	cggacgatgc	ggccgcgacc	tcgatcgatg	cgatggatgt	7860
agccggcctc	gtcgaagcag	cgctgggcga	acgcgagtcc	tgagaccgcc	gacctgggag	7920
atgacggtga	ccaccagtta	cgaagaagtt	gtcgaggcac	tgcgagcatc	gctcaaggag	7980
aacgaacgcc	tccggcgcgg	cagggatcgg	ttctccgcgg	agaaggacga	tcccatcgcg	8040
atcgtggcga	tgagttgtcg	ttatcccggg	cagggtctct	cgccggagga	cctgtggcaa	8100
ctgggtgccg	gcggtgtgga	cgcgatctcc	gaagtccggg	gggatcgccg	atgggacctg	8160
gatggcgtgt	tcgttccgga	ctccgatcgt	cctggcacgt	cgatagcctg	cgccggcggt	8220
tttcttcagg	gcgtgtcgga	gttcgacgcg	ggtttcttcg	ggatttcgcc	gcgtgaggcg	8280
ctggcgatgg	atccgcagca	gcggttgctg	ctggaagtcg	cgtgggaggt	cttcgagcgg	8340
gctgggctgg	agcagcggtc	gacacgcggg	tcccgcgttg	gcgtgttcgt	cggcaccaat	8400
ggccaggact	acgcgtcgtg	gttgcggacg	ccgccgcctg	cggtggcagg	tcagtgtgctg	8460
acgggcgggtg	cggcagcggt	tctttcgggc	cggttgctg	attcgttcgg	gttcgagggt	8520
cctgcgggtga	cggtggatac	ggcgtgttcg	tcgtcgttgg	tggcgttgca	cctggcgggg	8580
caagcactgc	gggccgggtga	gtgcgacctt	gcccttgccg	gtggcgtcac	ggtgatgtcg	8640
acgccgaagg	tggttctgga	gttctccccc	caacgggggtc	tcgcgcggga	tgggcgggtgc	8700
aagtcgttcg	cggcgggtgc	ggatggcact	ggatgggggtg	aggggtgccgg	actgttggtg	8760
ctggagcggt	tgtcggatgc	ccggcggaat	gggcacgagg	tgctggcggt	tgttcgtggt	8820
agtgcgggtga	atcaggacgg	tgcgtcgaat	ggtttgaccg	cgccgaatgg	ttcgtcgcag	8880
cagcgggtga	ttaccagggc	gttggcgagt	gcgggggtgt	cggtgtccga	tgtggatgct	8940
gtggaggcgc	atgggacggg	cacgcggctt	ggtgatccga	tcgaggcgca	ggcgctgac	9000
gccacctacg	gccgtgatcg	tgatcctggc	cgcccggtgt	ggttggggtc	ggtcaagtcg	9060
aacatcggtc	atacgcaagc	ggcggcgggg	gtggctggtg	tgatcaagat	ggtgatggcg	9120
atgcggcacg	ggcagctgcc	acgcacgttg	cacgtggaat	cgccgtcgcc	ggaggtggat	9180
tggtcggcgg	ggacggttca	actccttacg	gagaacacgc	cctggcccag	gagtggtcgt	9240
gttcgtcggg	tgggggtgtc	gtcgttcggg	atcagtggta	ctaacgcgca	cgtcacctc	9300
gaacagcccc	cgggagtgcc	gagtcagttc	gcggggccgg	gttcgggttc	tgtcgtggat	9360
gttccgggtg	tgccgtggat	ggtgtcgggc	aaaacacccg	aagcgctatc	cgcgcaggca	9420
acggcggtga	tgacctatct	ggacgagcga	cctgatgtct	cctcgcgtga	tgttggttac	9480
tcgttggcgt	tgacacggtc	ggcgttggtg	gagcagcgcg	tggtgctggg	gtcggaccgt	9540
<del>gaaaegttgt</del>	<del>tgtgcgggtg</del>	<del>gaaagcgctg</del>	<del>tctgccgggtc</del>	<del>atgaggcttc</del>	<del>tgggttggtg</del>	<del>9600</del>
accggatctg	tgggggctgg	gggccgcac	gggtttgtgt	tttccggtca	gggtggtcag	9660
tggctgggga	tgggcccggg	gctttaccgg	gcttttcggg	tggtcgtgc	tgcccttgac	9720
gaagcttgtg	ccgagctgga	tgacacatct	ggccaggaaa	tcgggggttcg	ggaggtggtg	9780
tccggttcgg	atgcgcagtt	gctggatcgg	acgttggtgg	cgcagtcggg	tttggttcgg	9840
ttgcaggtgg	gcttgctgaa	gttgctggat	tcgtgggggg	ttcggccgag	tgtggtggtg	9900
gggcattcgg	tgggcgagtt	ggcggcgggc	ttcgcggcgg	gtgtggtgtc	gttgctcggg	9960

gcggctcggt tgggtggc tegtgcccg ttgatgcagg cgttgccgc tggcgggtggg 10020  
 atgctggcgg tgccctgctgg tgaggagctg ttgtgggtcgt tgttgccga tcagggtgat 10080  
 cgtgtgggga tcgccgcggt caacgctgcg gggtcgggtg tgctctctgg tgatcgggat 10140  
 gtgctcgatg accttgccgg tcggctggac gggcaaggga tccggtcgag gtggttgccg 10200  
 gtgtcgcatg cgtttcattc gtatcggatg gatccgatgc tggcggagtt cggcgaattg 10260  
 gcacgaaccg tggattaccg gcgttggtgaa gtgccgatcg tgcgcacctt gaccggagac 10320  
 ctcgatgacg ctggcaggat gagcggggcc gactactggg tgcgtcaggt gcgagagccg 10380  
 gtccgcttcg ccgacgggtg ccaggcgctg gtcgagcacg atgtggccac tgttgctcag 10440  
 ctcggtccgg acggggcggt gtcggcgctg atccaggaat gtgtcgccgc atccgatcac 10500  
 gccgggcccg tgagcgcggt cccggcgatg cgcaggaacc aggacgaggc gcagaagggtg 10560  
 atgacggccc tggcacacgt ccacgtacgt ggtggtgcg tggactggcg gtcgttcttc 10620  
 gccggtacgg gagcgaaaca aatcgagctg cccacctacg ccttccaacg acagcgggtac 10680  
 tggctggtgc catcggaattc cggatgatgtg acagggtgccg gtctggcccg ggcggagcat 10740  
 ccgctgttgg gtgctgtggt gccggtcgcg ggtggtgacg aggtgttgct gaccggcagg 10800  
 atttcgggtg ggacgcattc gtggctggcc gaacaccggg tgctgggtga agtgatcgtt 10860  
 gcgggcaccg cgttgctgga gatgccttg cagcggggg aacgtcttgg ttgtgaacgg 10920  
 gtggaagagc tcaccctgga agcacgcgtg gtccctgccg agcgcggggc gatccaggtt -10980  
 cagctgcgag tgggcgcgcc cgagaattcc ggacgcaggc cgatggcgct gtattcacgc 11040  
 cccgaagggg cggcggagca tgactggacg cggcacgcca cgggcccgtt ggcgccaggc 11100  
 cgcggcgagg cggctggaga cctggccgac tggccggctc ctggcgcgct gccggtcgac 11160  
 ctcgacgaat tctatcggga cctcgcagag cttgggctgg agtacggccc gatcttccaa 11220  
 gggctcaagg cggcctggcg gcaaggggac gaggtgtacg ccgaagccgc gctgccggga 11280  
 acggaagatt ctggtttcgg ggtgcatccg gcaactgctg acgcggctct gcacgcaacg 11340  
 gctgtccgag acatggatga cgcacgcttg ccgttccagt gggaagggtg gtccctgcac 11400  
 gccaaaggcc cgcgggcttt gcgggtccgc gtggtcccgg ctggtgacga tgccaagtcc 11460  
 ctgctggttt gtgatggcac cggtcgaccg gtgatctcgg tggaccgact cgtattgcgg 11520  
 tcggctgcgg ccggcgagc cggtcgcgc cgacaggccc atcaagctcg gttgtaccgg 11580  
 ttgagctggc caacggttca actgccgaca tccgctcagc caccgtcctg cgtgcttctc 11640  
 ggcacctcag aagtgtccgc tgacatacag gtgtatccgg acctccggtc gttgacggct 11700  
 gcgttgatg ccggtgccga accaccggc gtcgtcatcg caccacgcc cccggcggt 11760  
 ggacgaacag cggatgtccg ggagacgact cggcatgcac tcgacctggt acaaggctgg 11820  
 ctttccgatc agcgactcaa cgaatccga ttgctcctgg tgacacaggg agcagtggcc 11880  
 gtggagccgg gcgaaccgt gaccgatctg gcgcaggccg cgctctgggg actgctgcgg 11940  
 tcgacgcaga ccgaacaccc tgatcgcttc gtcctcgtcg atgtgcctga gcccgcgcaa 12000  
 ctctccccg cgtgcccgg ggtgctggc tgcggcgaa ctcagctcgc gttgcgacgt 12060  
 ggcggcgctc atgcgccag actggctgga ctgggcagcg atgacgtcct gcccgtgccg 12120  
 gacggcaccg ggtggcgatt ggaggccacg cggccgggaa gcctggatgg gttggcattg 12180  
 gtggacgaac cgacggccac ggcaccgctg ggtgacgggt aggtcaggat tgcgatgcgc 12240  
 gcggccgggg tgaacttccg ggatgcgctc atcgcgctcg gtatgtatcc cgggtgtggca 12300  
 tcgctgggca gtgagggcgc cggggtcgtg gtggagaccg gcccggcggt caccggcctg 12360  
 gcaccggcg accgctgat gggaatgac ccgaaggcgt tcgggcccgt cgcggtcgcc 12420  
 gaccatcgca tggtagcag gattcccgt ggttggagct tegegegge cgcategggt 12480  
 ccgatcgtct ttctaccgc ctactacgc ctggttgatc tcgccgggtt gagaccaggg 12540  
 gagtcgttgc tggttcattc ggccgcccgt ggggtgggga tggccgcgat ccaactcgcc 12600  
 aggcacctg gtgcagaggt gtacgccacc gtagcgagg acaagtggca agccgtggag 12660  
 ctgagccgag aacacctgc ttcgtcgcgg acgtgcgatt tcgagcagca gttcctcggg 12720  
 gcaaccggcg gacgcggcgt cgacgtcgtg ctcaactccc tcgccgggga gttcgccgat 12780  
 gcgtctctgc gaatgctgcc gcgcggtggc cgtttcctgg agttggggaa gacggatggt 12840

cgtgaccccg	tcgaggtcgc	cgatgcgcat	cggggcgtgt	cttaccaggc	tttcgatacc	12900
gtagaggcag	gcccgcagcg	aatcggcgag	atgcttcacg	agctggtgga	gttggttcgag	12960
ggacgcgtgc	tggagcccct	gcctgtcacg	gcttgggacg	ttcggcaggc	gcccgcagcg	13020
ctacggcacc	tgagccaagc	gcggcatgtg	ggaaagctgg	tgctcaccat	gcctccggtg	13080
tgggacgccc	caggcacggt	tctggttacc	ggcggaaacg	gagcacttgg	cgcagaggtc	13140
gcccggcacc	tcgtgatcga	gcgcgggggtg	cgaaacctgg	tcctcgtcag	caggcgcggt	13200
cccgcagcca	gtggcgctgc	tgagctcgtg	gcgcaactga	cggcctacgg	tgccgaggtt	13260
tccttgacag	cttgcgatgt	cgccgatcgt	gagaccttgg	cgaaggtgct	tgccagcatc	13320
ccggacgagc	atccgttgac	cgccgtgggtg	cacgcggctg	gtgttctcga	cgacggagtg	13380
tccgaatcgc	tcaccgtgga	gcggctggac	caggttctgc	gcccgaaggt	cgatggcgcg	13440
cggaatctgc	tcgagctgat	cgaccgggac	gtggccctcg	tggtgttctc	gtcgggtgctg	13500
ggtgtgctcg	gcagcggtgg	gcagggtaac	tacgcggcgg	ccaactcctt	cctcgacgca	13560
ttggcgacgc	aaaggcagtc	gcgcggccta	ccgacgagat	cattggcctg	ggggccctgg	13620
gcggaacatg	gcatggccag	caccttgccg	gaagccgagc	aggatcgatt	ggcgcgatct	13680
gggttgctgc	cgatctcgac	cgaggagggg	ttgtcccagt	tcgacgccgc	gtgcggcggc	13740
gcgcataacc	tggtagcgcc	ggttcgattc	agccgcttgt	ccgacgggaa	cgcgatcaag	13800
ttctccgtcc	tgcaaggttt	ggtcggggcg	catcgcgctca	acaaagcggc	gactgcggat	13860
gatgccgaga	gcctccggaa	acggttggga	cgcttgccgg	atgcagaaca	acatcggatt	13920
ctgctggacc	tcgtccgcat	gcatgtggcg	gcagtgtctg	gattcgccgg	ttctcaggag	13980
atcacccggg	acggcacggt	caaggtgctg	ggcttcgact	cgttgaccgt	ggtcgagttg	14040
cgcaaccgga	tcaacggggc	gacggggctg	cgactgcccg	ccaccctggt	gttcaactac	14100
ccgacgccgg	atgcgctcgc	cgcgcacctc	gtcacccgcg	tgtccgcaga	ccgcctggcc	14160
gggacattcg	aggaactcga	caggtggggc	gcgaacctgc	ccacgctggc	cagggatgag	14220
gccacgcggg	cgcagatcac	caccgggcta	caggcgatct	tgacagacct	ggcgacgtg	14280
tccggcgga	ccggcgggcg	ctccgtgccg	gaccggctca	gatcgggcac	ggacgacgag	14340
cttttccaac	tcctcgacaa	cgatctcgaa	cttccctgat	gcctcagccg	gagccttcgc	14400
aacttccctg	agggaaacgc	cacatgtcga	atgaagagaa	gctccgggag	tacttgccgc	14460
gtgcgctcgt	ggatctgcac	caggcgcgcg	agcggctgca	cgaggcggag	tcgggagagc	14520
gggaacccat	cgcgatcgtg	gcgatgggct	gccgggtacc	gggtgggggtg	caggaccctg	14580
aagggtctgt	gaaactggtc	gcctccggtg	gcgacgccat	cgggtgaattc	cccgtgatc	14640
gtgggtggca	cctcgacgag	ctctacgac	ccgaccggga	tcagcccggga	acctgctaca	14700
cccggcacgg	cggcttcctc	cacgacgccg	gcgagttcga	cgcgggattc	ttcgacatca	14760
gccccctgta	ggcgctcgcg	atggaccgcg	agcagcggct	gctgctggaa	atctcctggg	14820
agaccgtcga	atccgctggg	atggaccgga	ggtccttgcg	ggggagccgc	accggggtgt	14880
tcgcgggatt	gatgtacgag	ggctatgaca	ccggcgccca	ccgggcagga	gaaggtgtcg	14940
aaggctatct	cggaaaccggc	aatgcgggaa	gcgtcgccctc	tggtcggggtt	gcgtatgcgt	15000
tcgggttcga	gggcccagcg	gtgacggtag	acacggcgctg	ctcgctcgtcg	ttgggtggcg	15060
tgcatattgg	gtgtcagtcg	ttgcggcagg	gcgagtgtga	tctggcgctg	gccggtggag	15120
tgacggtgat	gtcgacgccg	gagaggttcg	tggagttctc	ccgtcagcgt	ggtctcgcac	15180
cggatgggcg	gtgtaagtcg	ttcgcgggcg	ctgcggatgg	aaccggtttg	ggtgaggggtg	15240
ccggtttggt	gttgctggag	cggctgtcag	acgccaggcg	gaacgggcat	cgggtactgg	15300
<del>eggttggtcg</del>	<del>tggtagcgcg</del>	<del>gtgaatcagg</del>	<del>acggtgctgc</del>	<del>gaacgggattg</del>	<del>acggggccga</del>	<del>15360</del>
acgggctggc	ccaggagcgg	gtcattcagc	aggtgctcac	gagtgcgggg	ctgtcggcgt	15420
ccgatgtgga	cgtgtggag	gcgcattgga	cgggtacgcg	gcttggtgat	ccgatcgagg	15480
cgcaggctct	gatagccgcc	tatggacagg	atcgggaccg	ggaccggccg	ctgtggttgg	15540
ggtcgggtcaa	gtccaacatc	ggtcatacgc	aggcggctgc	gggcgtcgct	ggtgtgatca	15600
agatggtcat	ggcgatgcgg	cacggggagc	tgccgcgcac	gttgcacgtg	gacgagccga	15660
attcgcacgt	ggactggtcg	gctggtgcgg	tccgactcct	gaccgagaac	atccgctggc	15720

cagggacggg tacgcg gctggagtgt cgtcgttcgg ggtaaa gt accaacgcac 15780  
acgtcatcct cgaacacgac ccgctcgcgg tgaccgagaa cgaggaagca gcgcagtcac 15840  
cagcacctgg gatcgtgccc tgggcgttgt ccggggcggtc gtgcacggcg ctggcgggccc 15900  
aggccgaacg gctgcgcgag ctgtgcgagc agaccgatcc cgaccccgtc gatgtcggtt 15960  
tctcactggc cgccacgcgc acggcttggg agcaccgagc ggtggtgctt ggtcgggaca 16020  
gcgctacgtt gcgctccggg cttggcgctt ttgccagcgg tgaaccagcg gtcgatgtcg 16080  
ttgaggggag cgtcctggac ggcgaggtcg tcttcgtctt ccccggtcag ggctggcagt 16140  
gggcccgtat ggcagtcgac ctgctggacg cttcgccgac gttegcgcgc cacatggacg 16200  
agtgcgccac cgcgctgcgg aggtacgtgg actggtcgtt ggtcgacgtg ctgcgcggag 16260  
cggagaactc cccaccgctg gaccgggtgg acgtgctcca gcccgctcc ttcgcggtga 16320  
tgggtgtcgt cgccgaggtg tggcgcttct acggggtgag gccggcgcc gtcgtcgcc 16380  
acagtcaagg cgaaatcgcc gcggcctgcg cagccggggt gctgccgctg gaggatgcgg 16440  
ccaggcttgt cgcattgcgc agcagagcgt tgaagggact ttcggggcgg ggtggcatgg 16500  
cgtcgttggc ctgcctgcg gatgaggtcg cggcattgtt cgcgggatcg ggcggccgctc 16560  
tggaagtgtg ggcgatcaac ggcccgcat cggctcgtgt gtccggcgat ctggaagcgg 16620  
tggaagaaact gctggcagag tgcgtgaaa aggacatgcg tgcacgccgt atccccgtcg 16680  
actacgcctc gcattcagcg cacgtggagg tggctcggag ccgggtgctg gcggccgccc 16740  
ccggggtgcg acaccgggac ggccaggtgc cgtggtggtc gacggtgatc ggcgactggg 16800  
tgatccggc caggctggac ggcgagtatt ggtatcgaa cctccggcag ccggtccggt 16860  
tcgaacacgc cgtgcagggc ctggtcgagc ggggattcgg cctgttcac gaaatgagt 16920  
cgcacccggt gctgaccacg gcggtcgagg aaaccgggtc ggagtcggag accgccgtgg 16980  
ccgcggtagg taccttgcca cgtgactcgg gcggcctccg gaggttgttg cattcgctgg 17040  
ccgaggcgta cgtgcgcggc gccaccgtgg actggggcgt ggcttcggg ggcgcggggc 17100  
gacggctgga cctgccgacc taccggttcc agcgccagcg gtactggctg gacaaggag 17160  
ctgctccga cgaggctcgt gcggtctcgg acccgcgccg gggctggtt tggcaagccg 17220  
tggcgcgcca agacctgaaa agcgtgtccg atgccctcga tctcgacgcc gacgcaccgc 17280  
tgagcgcaac acttccagcc ctgtcgtctt ggcaccgtca ggaacgagaa agggcttgg 17340  
cagacggttg gcggtaccga gtgcactggg tacgggtggc cccgcagccg gtccggagaa 17400  
cgcgggaaac ctggctcctg gtggttcccc cgggcggcat cgaggaagcg ctggtcgaac 17460  
ggctgacgga tgcgttgaa acgcgagga tcagcacct gcgcctcgac gtgccaccgg 17520  
cggcgaccag tggcgaactc gcaaccgaac tccgcgcgc agccgacgtt gaccgggtga 17580  
aggcaatcct gtcgctcacc gcgttgagc agcgaccca cccgaatgc aaggacgtcc 17640  
cgagcgggat tgccttgctg ctgaacctgg tcaaggcgct cggtaagcc gacctcagaa 17700  
ttcctctgtg gaccatcacg cgtggtgcgg tcaaggcagg ccccgagat cggctgctgc 17760  
gcccgatgca ggcgcaagca tggggtctgg ggcgagtagc cgactcgaa cccccgagc 17820  
gctggggttg gctgatcgac ctgccggtt cgtggacgg cgacgtcctc acgaggctgg 17880  
gcgaagcgct caccaacggc ttggcggaag accaactggc gattcgccag tcggcgctgc 17940  
tggcccgcg actggtaccc gccccggcga atcagccgc tggacgtaag tggcgcccc 18000  
gagggagcgc gctgatcacg ggcggactcg gcgcgggtgg cgacaggtg gcgaggtgg 18060  
tggccgaaat cggagccgag cgaatcgtgc tcaccagtc acggggcaac caagcagcag 18120  
gcgcgcgcca gctggaagcc gaactccggg ccttgaggc gcaagtgtcc atcgtggctt 18180  
gcgacgtgac cgatcgtgcc gagatgtccg cactactggc cgagttcgac gtcaccgcg 18240  
tgttccacgc ggcggagtc ggtcggctgc tgccgttggc ggagaccgac cagaacggcc 18300  
tggccgaaat atgcgcggcg aaggctccgc gcgctcaggt gctggacgaa ctgtgcgaca 18360  
gcaccgatct cgatgccttc gtctgttct cctcgggtgc cggggtatgg ggcggggcg 18420  
gtcagggcgc ttacggcgcg gcgaacgcat tcttgacac actcgccgaa caacgccgag 18480  
cacgcggtct gccggcaacc tcgatctcct ggggcagttg ggcggcgcc ggcatggccg 18540  
acggcgcgcc gggcgaacac ctgcggcgac gcgggatacg tccgatgccg gcggcgctcg 18600

ccatcctggc tctgcaggaa gtacttgacc aggatgagac gtgcgtgtcg atcgctgatg 18660  
tggactggga ccgattcggt cccacgttcg ccgcgactcg cgccaccggg ttgttcgacg 18720  
aagtgcgggc ggcgagaaaag gcgatgcccg cgaatggggc ggcagaacca ggcggctcgc 18780  
cgttcgcccg caatctcgcg gagctgccgg aagcccaacg acgccacgaa ctggtggatc 18840  
tgggtgtgcg ccaggtggca accgtgctcg ggcacggcag tcgcgaggaa gtccagcccg 18900  
agcgggcggt ccgcgcgctc gggttcgact ccctcatggc ggtggatctg cgcaatcggt 18960  
tgaccaccgc caccgggttg cgcctgccga ccacaaccgt cttcgactac ccgaatccgg 19020  
ccgccttggc cgctcacctg ctcgaggagc tgggtgggtga tgtcgcgtcg gctgcgggtga 19080  
ccgctgccag cgcgcccgcg agtgacgaac cgatcgcgat cgtcgcgatg agctgccggg 19140  
ttccgggtgg cgcgcactcg ccggaagacc tgtggcggct ggtcgcggcc ggcacggagg 19200  
tgatcggcga gttccctcc gaccggggct gggatgcgga aggcctttac gatccggatg 19260  
cttcaggcc tggaacgacg tatgcgcgga tggcgggatt cctctacgac gccggtgagt 19320  
tcgatgccga cctgttcggc atcagcccac gtgaggcggt ggcgatggat ccgcagcagc 19380  
ggttggtgct cgaaatcgcc tgggaagccc tcgaacgggc cggaatcgat ccgttgtcct 19440  
tgaagggcag tggggtcggc acgtacatcg gcgctggaag ccgtgggtac gcgacggatg 19500  
tgccggcagtt tcccgaggag gcgaggggct acctgctgac gggtagctcg gccagtgtgc 19560  
tgtcgggtcg ggtcgcgtat tcgtttgggt tcgagggtcc tgcggtgacg gtggatacgg 19620  
cttggtcgtc gtcgttggtg gcgttgcatc tggcgtgccg gtcgttgctg tcgggcgagt 19680  
gtgatctggc gttggccggg ggtgtgaccg tgatgtcgac gccggagatg ttcgtggagt 19740  
tctcccgta gcgcggtttg gcgccggatg ggcgggtgca gtcgttcgcg gagagcgcgg 19800  
acggcaccgg ctggggcgaa ggcgcggggc tgttgttgct ggagcggttg tcggacgccc 19860  
accggaatgg gcatcgggtg ttggcgggtg ttcgtgggtc agcggtgaat caggacggcg 19920  
cctcgaacgg actggcggcg ccgaacggtc cgtcgcagca gcgggtgatc aaccaggcac 19980  
tcgcgaatgc ggctctttcg gcgtccgatg tggatgcggg ggaggcacat ggcaccggga 20040  
ccaggctggg tgatccgatc gaggcgcagg cattgatcgc aacgtatggg caggcccggg 20100  
agcgggatcg gcccttgtgg ctggggtcgg tcaagtcgaa catcggtcat acgcaggccg 20160  
cggcgggtgt tgccgggtgt atcaagatgg tgatggccat gcggcacggg cagctgcccg 20220  
cctcgctgca cgcggatgag cccacgtcgg aggtcgattg gtcgtcgggg gcggtccggc 20280  
tctcgcgga acaggtaact tggccggagt ctgaccgtgt tcgtcgggtg ggggtttcgt 20340  
cgttcgggat cagcggcacc aacgcacatg tgatcctcga acaagctacg aatgcgccag 20400  
atagtacagc ggagacggac aaaacagaat ccgatctac tgtcgatatt ccggtcgttc 20460  
cctggttggt gtcgggaaag acgacggatt ccctgcgggg acaagccgaa cgagtcttgt 20520  
ctcaggctga gtcccggccg gagcagcggt cgctggatgt tgcctactcg cttgcttctg 20580  
gccgagccgc gctggatgaa cgcgctgtcg tgctgggtgc ggaccgcggg gagctggttg 20640  
ctggactggc ggcgttggcc gccggtcagg aggtctctgg ggtgatcagc ggaactcgtg 20700  
cttctgctcg gttcgggttc gtgttctcgg ggcagggtgg tcagtgggtg gggatgggca 20760  
gagcgtcta ctcaagttt cgggtgttcg ctgctgcgtt tgatgaggct tgcgccgagt 20820  
tggaggcaca tctgggggaa gaccgccggg ttcgggatgt ggtcttcggg tccgatgcgc 20880  
agctgctgga tcagacgctg tgggcgcagt cgggtctgtt cgcgctgcaa gccggcctct 20940  
tggggctgct gggttcgtgg ggcgttcggc cgatgtggg gatggggcat tcggtcgggg 21000  
agttggccgc cgcgtttgcg gctggcgtgt tgcgttgctg ggatgcggct cggttggtgg 21060  
cegegegege ceggttgatg caagccctgc cctctgacgg cegatggtg geggtggetg 21120  
ctggtgaaga ccttgttcgg ccattgctgg ccggtcggga ggagtcctg agcgtgcccg 21180  
cgctcaatgc ccccggttcg gtggtgttgt cgggcgatcg ggaggtgctg gccagcatcg 21240  
tcggccgggt gaccgagctc cgagtccgga cgcggcgctt gcgggtctcc catgcttttc 21300  
attcgacccg gatggacccg atgttgggag agttcgccca gatcgccgag tctgcggagt 21360  
tcggtaagcc aacgacaccg cttgtgtcga cgttgacggg tgagctcgac agagccgcgg 21420  
aaatgagcac accagggtat tgggtgcgcc aggcgcgtga acccgctcgt ttcgccgacg 21480

gtgtccaggc cctggcagg cagggcatag gcacggctgt cgagctggc ccggacggaa 21540  
 cgctggcggc actggttcgg gagtggtcga ccgagtcga tcgggttggg cggatttcgt 21600  
 cgatcccact gatgcgcagg gagcgggacg agacccttc ggtgatgaca gccctggcgc 21660  
 atctccacac ccgtggtggt gaggtggact ggcaggcgtt ttctgccggt accggcgcta 21720  
 ggcagctcga gttgccaacg tatgccttc aacgacagca ctactggatc gactccagt 21780  
 cgcgccagc acgcgaccgc gcagacatcg gcgaggtggc ggaacagttc tggaccgcgg 21840  
 ttgaccaagg cgatctggca acgttggtcg ccgctctgga tcttggggcg gacgacgaca 21900  
 catgcgcatc gttgagcgat gtattgccgg cgttgctctc ctggcgaagc ggactccgca 21960  
 accgttcgtc cgtcgattcc tgccgggtacc gaatcagttg gcattcctct cgggaggtgc 22020  
 cggccccgaa gatttcgggt acctggctgt tggctcgtgc cgggtcgtcg gatgacggat 22080  
 tggtcacggc tttgacgagt tcaactggtc gaggcggcgc cgaggtcgtc cggatcggcc 22140  
 tgtccgaaga ggaccgcac cgcgaggacg tcgcacagcg gctggccaat gcgctgacgg 22200  
 atgccgggtca actcgggtggc gtgctttcgc tgttggggct cgatgaatcg cctgctccgg 22260  
 gattctcctg cttgccaact ggtttcgcgc tgactgtgca gcttctgcgg gccttgccga 22320  
 aggcgcagct cgaggcgct ttttgggcgg tgacgcgcgg cggcgtcgcg ttggaagatg 22380  
 tacgcgtgtc tccggagcag gccctggtct gggggtgct gcgtgtcgcg ggactggagc 22440  
 acccgaggt ctgggggtggc ttgatcgacc tgccatcgga ctgggacgac cgattgggtg 22500  
 cccggttggc ggggtgtgtg gcggtggtg gcgaggatca agtcgccatt cgcgtgggtg 22560  
 gtgtgttcgt gcggcggttg gaacgcgctg gtgcgtcggg tgccgggtcg gtgtggcgtc 22620  
 ctccgggggac ggtgttgggt acgggtggtg cgggcggtt gggggcgcat gttgcccggt 22680  
 ggttggccgg tgccggggct gagcacgtgg tgttgaccag ccgtcgagga gcggacgctc 22740  
 cgggcgctgg ggaattgcgg gcggagctgg aggcgctggg tgctcgggtg tcgattgtgc 22800  
 cctgcgacgt ggctgatcgt gacgcagtg ctggagtgtt ggcagggatc ggtggggagt 22860  
 gtccgctgac tgccggtgta cacgccgccg gggtcggcga ggcgggcgac gtagtgga 22920  
 tgggtttggc ggattttgca gcggtgtgt cgcgcaagg gcgtgggtgc gcgaatctgg 22980  
 acgagttgct ggccgactcg gagctggatg cgtttgtgat gttctcctcg gtgtcggggg 23040  
 tgtggggagc cggcgacag ggtgcgtatg cggctgcgaa cgcctacttg gatgcgttg 23100  
 ccgagcagcg tcgggcgagg ggattggtcg ggaccgcggt tgctggggga ccgtgggccc 23160  
 gtgacggcat ggccgcggc gaaaccggcg cacagctgca ccggatgggc ctggcgtcga 23220  
 tggaaaccgag cgccgcgctg ctggcacttc aggtgcat ggaccgcgat gagacctccc 23280  
 tcgtcgtggc cgatgtcgat tgggcacggt tcgccccagc cttcacctcg gcacgtcgac 23340  
 gcccgctgct ggacaccatc gacgaggccc gagccgcat ggaaaccacc ggcgaacaag 23400  
 cgggcacagg caaaccggt gagctgacgc aacgcctggc cggactgtcg cggaggaac 23460  
 gcgacgatgc ggtattggat ctggtgcggg cggagacggc ggctgtgctg ggacgcgacg 23520  
 atgccacggc cctggcgcca tcgcgccgt tccaggaact cggattcgac tcttgatgg 23580  
 cgggtggagct gcgcaaccgg ctgaacaccg ccaccgggat ccagctgcc gccagcagca 23640  
 ttttcgacta cccaatgcc gagtcgctgt cgcgtcacct ctgcgccgag cttttcccaa 23700  
 cggagactac cgtggactcg gcccttgccg agctcgatcg aatcgagcag cagctctcga 23760  
 tgctcaccgg cgaagcgcg gcacgggacc gaatcgcgac acgactgca gccctccacg 23820  
 agaagtggaa cagcgcagct gaagtaccga ccggagccga tgcctgagc acgctcgatt 23880  
 cggcgacgca cgacgagata ttcgagttca tcgacaacga gctcgacctg tctgagcag 23940  
 ttcctgggga acttcaggc cggaaatcg gtggaaatca caatggccaa tgaagaaaag 24000  
 ctcttcggct atctgaagaa ggtaactgcg gacctgcatc agaccgggca gcgcctgctc 24060  
 gcggccgaga gccggagtca ggagccgatc gcgatcgtct cggcgagctg ccgactgcc 24120  
 ggcggcgctc actctcccga agcgtcttg caactcgtgc gactggcac cgacgccatc 24180  
 tcggagttcc ccgcgaccg gggctgggat ctccggcggg tgtacgatcc cgaccgaac 24240  
 caccaggga cgctgtacac gcgggcccgc ggtttcctcg caggagcggg cgatttcgac 24300  
 cccgccatgt tcgggatttc gcccgctgag gcgttggcga tggaccgca gcaacgggtt 24360



ttgctggagc	tgctcctggga	ggccctcgaa	cgggcgggca	tagacccgac	atccctgogc	24420
ggcagcaaga	ccggtgtctt	cgggtggtgc	acgccccagg	agtacggggc	gtccttgacg	24480
gagatgagcc	gaaacgctgg	gggttttgga	ctcaccgggc	ggatggtgag	tgtggcgctc	24540
ggtcggggtg	cgtattcggt	tgggttttag	ggtcctgcgg	tgacggtgga	tacggcggtg	24600
tcgtcgtcgt	tgggtggccct	gcatttgggc	tgtcagtcgt	tcggttccgg	cgaatgcgat	24660
ctcgcgctgg	ccggcgggtg	gacggtgatg	gcgacaccgg	cgacgttcgt	ggagttctcc	24720
cgtcagcgtg	gtttgggtcc	ggacggggcg	tgcaagtcgt	tcgcggtctc	cgcggtatgg	24780
accgggtggg	gtgaggggtg	cggctctggg	ttgctggagc	ggttgtcgga	tgcgcgggcg	24840
aatgggcacg	aggttctggc	ggtggtgcgg	ggtagcgcg	tgaaccagga	cggcgcgtcg	24900
aatggtttga	ctgcgcgcaa	tgggtccgtc	cagcagcggg	tgatcaccca	ggcgttggcg	24960
agtgcggggc	tgctcggtttc	cgatgtggat	gcggtcgagg	cacatgggac	cgggaccacg	25020
ttgggtgatc	cgatcgaggc	acaggccctg	atcgccacgt	acgggcaggg	ccgggagaag	25080
gatcggccgt	tgtggttggg	gtcgggtcaag	tccaacatcg	gtcacacgca	ggcgcccgct	25140
ggcgttgccg	gcgtcatcaa	gatggtcttg	gcgatgcggc	acgggcagct	gcccgccacg	25200
ttgcatgtgg	atgagccac	gtcggcggtg	gactggtcgg	cgggttcggg	ccggcttctc	25260
acggagaaca	cgccctggcc	ggacagtggg	cgtccttgcc	gggtgggggt	gtcgtcgttc	25320
gggatcagcg	gcaccaacgc	acatgtgatt	ctcgaacagt	ctccagtcga	gcagggcgaa	25380
ccggccgggc	cggtcgaagg	cgagcgggaa	ccggatgtag	ccgtccccgt	ggtgccttgg	25440
gtgctgtcgg	gtaagacacc	ggaggctgcg	cgggcgcagg	ccgaacgggt	gcattcgcat	25500
atcgaggacc	ggccgggggt	gtcgccgggt	gatgtggcgt	attcgctagg	aatgacacgc	25560
gcggcgctgg	atgaacgcgc	agtgggtgtg	ggctcggacc	gtgccgcgct	cctgaccggg	25620
ttgagggcat	tcgccgacgg	ctgcgatgcg	cccgaagtgg	tttcgggggt	tgtggggctt	25680
ggtggcccg	tcgggttcgt	gttctcgggt	cagggtgggt	agtggccggg	gatgggccc	25740
gggctctact	cgggtgttcc	ggtgttcgcc	gacgcgttcg	acgaggcttg	cgcgagttg	25800
gatgcacacc	tgggccagga	actgcgggtt	cgggatgtgg	tgttcggttc	gcaagcgtgg	25860
ttgctggatc	ggacgggtgt	ggcgacgtcg	ggtttgttcg	cgttgcagat	tggcttgctg	25920
cggctgctgg	gttcgtgggg	tggttcggcc	gatgtggtgt	tggggcactc	ggtgggtgag	25980
ctggctgcgg	tgcatgcggc	tggtgtgttg	tcgttgctcg	aggccgcgcg	gttgggtggc	26040
ggtcgcgccc	ggttgatgca	ggcgttgcc	tctggtggtg	ccatgctcgc	ggtcgctacg	26100
ggtgagtttc	aggtcgatcc	tctgctggat	ggggtgcggg	accggatcgg	tatcgcgcg	26160
gtgaatggcc	cggaatcggt	tgtgctctct	ggtgaccgcg	agctgctcac	cgagatcgct	26220
gatcggttgc	acgatcaggg	gtgccggacc	cgggtggttc	gggtgtcgca	tgttttccat	26280
tcgccccata	tggagccgat	gctggaggag	ttcgcccaga	tctcccagg	ccgcgaatat	26340
cacgcaccgg	aactgccgat	catctcgacc	ctgatcggtg	agctggacgg	tggtcgagtg	26400
atgggcactc	ccgagtactg	ggtgcgtcag	gtgcgtgagc	ccgtccgttt	cgccgagggg	26460
gtccaggcgc	ttgtcgggtc	gggtgtcggc	acgattgtcg	aattgggtcc	ggacggggcg	26520
ttgtcgacgt	tggtcgagga	gtgtgtggcg	gaatccgggc	gggtggccgg	gatcccgtcg	26580
atgcgcaagg	accgcgacga	ggcgcgaaac	gtgctggcag	ctttggcgca	gatccacacc	26640
cgtggtggtg	aggtggactg	gcggtcgttt	ttcgccggta	ccggggcgaa	gcaagtcgac	26700
ctgccacact	acgccttcca	gcggcagcgg	tactggtggt	catccaccgg	gcgtgcgggt	26760
gacgtgaccg	ccgcgcgatt	ggccgaggcg	gaccatccgc	tgctcgggtg	ggtggttgcg	26820
ttggcagacg	gcgaaggtgt	ggtgctgacc	ggtcgggtga	cagcgggttc	gcattccgtg	26880
ttgtccgatc	accgggtgct	gggcgaaatc	gtcgtccccg	gcaccgcgat	cgtcgagctg	26940
gtgtggcacg	tcggcgagcg	cctcggttgt	ggccgggtgg	aagaactggc	tttggaagcg	27000
cccctgatcc	tgccggatca	tggagcggtc	caggttcagg	tgtggtggg	accgcccggg	27060
gaatccggag	ccgggtcggg	ggcgctctac	tctgtctctg	gcgaggcgat	cgaacccgag	27120
tggaagaagc	acgcgacggg	cgtgcttctc	ccaccggtgg	ccgcgcgagaa	ccatgagctg	27180
accgcatggc	ccccggagaa	tgcgaccgaa	atcgatgcag	acggggtcta	cgcattcctt	27240

gaagggcacg gtttcgca cggaccggcc ttagatgtc tgcgcggtgc ctggcgacga 27300  
 ggcggggagg tgttcgccga agtcgcattg ccgatgaca tgcaggcggg ggtcgatcga 27360  
 ttcggcgccc accccgcgtt gctggacgcg gttctgcatg ccgccgcagc cgagacgtcg 27420  
 gtgggtccaga gcgaagcgcg ggtgccgttc tcgtggcggtg ggggtggaact tcgcgccact 27480  
 gaaagcgcgg tgggtgcgggc gcgcctctcg ttgacttcgg atgacgaact gtcgttggtc 27540  
 gcagtggacc cggctggccg attcgtggcc acggttgatt cgctggtgac ccgaccgatc 27600  
 tcccggcagc aggtgaggtc tggcgcgatc ggtgattgcc tgttcgaggt ggagtggcac 27660  
 cggaaggcgt tgttggaac aaccgcggc gacgacctt ccatcgtcgg tgacggtccc 27720  
 agttggccgg aatcgggtgc cgcaaccgca cggttcgcca ccctggatga gttccgtgcg 27780  
 gccgtggact cggacgttcc tgccccgggt tcggtgttg tgcagctat gtcggccgaa 27840  
 gaggtcgagg gtgatccct gccgtcgcgc gcccaagagt cgacctcga tctgctggct 27900  
 ctcgtgcagt cgtggcttgc ggacgagcgg ttcgccgaat ccagctcgt ggtcgtcacg 27960  
 cgtgcagcgg tgcggccga ctcggattcg gacgtcgcgg acctggtggg tgcgtcgtcg 28020  
 tgggggttgt tgagttcagc ccagtcggag aaccgggtc gcttcgtgct ggtggacgtg 28080  
 gacggcacac ctgagtcgtg gcaggcgtt cggccgcgcg tgcgagcagg agaaccgcag 28140  
 ctggcacttc ggcgcggcgt ggcgtggtg cctcgggttg cgcgactcac ggtgcgcgag 28200  
 gagggtcct ccccgcaact cgacacggac gggaccgtcc tcatcacggg tggcaccggt 28260  
 gcgttggggg gagtggttgc ccgtcacctg gtggaggagc acgggattcg gcgtttggtg 28320  
 ttggcaggcc ggcgtggtg gaatgcgcct ggagtccacg agttggtgga tgagctggcg 28380  
 cgcgcggggc ccgtggttga ggtggtggct tgcgatgtgg ctgaccgcac cgatctggag 28440  
 cacgtgctgg ccgccattcc ggtcgactgg ccgtcgcggg ggatcgtgca taccgctggg 28500  
 gtgctggccg acggagtgat cgggtccttg tcggcggcgg atgtgggcac ggtgtttgcc 28560  
 ccgaaggatga cgggggcatg gcatctgcac gagttgacct gcgatctgga tctgtcgttc 28620  
 ttcgttcttt tctcttcctt ctccgggatt gcgggtgccg cagggcaggc caactacgcg 28680  
 gcggcgaaca cgttcctgga tgcatggcg cgttatcgcc gggcgcggtg gctgcctggg 28740  
 ttgtcgttgg cgtggggact gtgggcgcaa ccagcggta tgacgagtgg cttggacgcg 28800  
 gcgtcgttgg agcggttggc gcggacgggc atcgcagaac tttccacgga ggatggactc 28860  
 cgcctgttcg atgccgcgtt cgcgaaggac cgggcttgcg tcgttgccgc tcgattggac 28920  
 agggcgctgc tggtcgggaa cggacgatcg cacgcgattc cggcgctgtt gagcgcgttg 28980  
 gttcctgttc gcggcggtgt ggcgaggaaa acagccaatt ctcaggccgc ggatgaggac 29040  
 gcaactgttg gtttggtgcg ggagcacgtt tcggccgtgc tgggttattc ggtgcggtc 29100  
 gaggttgggg gcgaccgtgc tttccgtgat ctgggttttg attcgttgte tggcgtggag 29160  
 ttgcggaacc gccttgccg ggtgctgggg gtgcggttgc cggcgactgc ggtgttcgac 29220  
 tatccgacgc cgcgggcgct ggcgcgttct ctgcatcagg aactggcagg cgaggtcgcg 29280  
 tccacgtcga cgcgggtgac cagggcagcg agtgccgaag aggatcttgt tgcgattgtc 29340  
 gggatgggat gtcgttttcc ggggtggggtg tcgtcgcggc aggagctttg gcggctggtg 29400  
 gccggcggcg tggatgcggt ggctgggttc ccagacgatc gcggctggga tctcgcggcg 29460  
 ttgtacgatc ctgatcccga tcgtctcggg acctcgtagt tgtgtgaggg cgggtttctg 29520  
 cgggacgcgg cggagttcga tgctgacatg ttcggcatca gcccgctga ggcgttggcg 29580  
 atggatccgc agcagcggtt gctgctggag gtcgcctggg aaaccttgga gcgggctggg 29640  
 atcgatccgt tctcgttgca cggcagccgg accggtgtgt tcgcgggctt gatgtaccac 29700  
~~gaetatgggg cccgatteat taacagagea cgggagggt tegaagggea cctcgggacg 29760~~  
 ggcaatgcgg ggagcgtgct gtcgggtcgg gttgcgtatt cgtttggttt cgagggtcct 29820  
 gcggtgacgg tggatacggc gtgttcgtcg tcgttggtgg cgttacacct ggcgggtcaa 29880  
 gcaactgcgg ccggtgagt cgaattcgcc cttgocgggtg gcgtcacggt gatgtcgacg 29940  
 ccgacgacgt tcgtggagtt ctcccgtaa cggggtctgg ctccggatgg gcggtgcaag 30000  
 tcgttcgcgg cggccgcgga tggcaccggg tggggcgagg gtgcgggtct ggtgttgctg 30060  
 gagcggttgt cggatgcccg gcgcaatggg cacgaggttc tggcggtggt gcggggtagc 30120

gcggtgaacc	aggacggcgc	gtcgaatggc	ttgactgcgc	caaatggtcc	gtcacagcaa	30180
agggatgatca	cccaggcact	cacgagtgcc	gggctgtccg	tgtccgacgt	ggatgctgtg	30240
gaggcgcacg	ggacgggcac	gcggcttggt	gatccgatcg	aggcgcaggc	ggtgatcgct	30300
acgtacggcc	gggatcgtga	tcccggctcg	ccgttgtggc	tggggtcggg	gaagtcgaat	30360
attggtcaca	cccaggcggc	ggcgggtgtc	gctggtgtga	tcaagatggt	gatggcgatg	30420
cggcaggggg	agctgccgcg	cacgttgcac	gtggacgagc	cctccgcgca	ggtggactgg	30480
tctgcgggca	cggtcacaact	cctcacggag	aacacgccct	ggcccgcacg	cggctcgtctt	30540
cgccggggcg	gcgtgtcatt	gttcggggatc	agtggcacca	acgcgcacct	gacccctgaa	30600
caacctccgc	gagatgcgca	gcgctcaaca	gagccggatt	cgggttctgt	ccgcgatttt	30660
ccggtggtgc	cgtggatggt	gtcgggcaaa	acaccgaag	cgctatccgc	ccaggcagat	30720
gcattgatgt	cctacttgag	caatcgcgtt	gatgcttccc	cgcgagatat	cggttattcg	30780
cttgcggtga	cccgtccggc	gttggaccac	cgcgctgtcg	tgctgggtgc	ggatcgtgcc	30840
gcgttgctgc	cgggcttgaa	agcgttgccc	gttagtaatg	acgctgccga	ggtgatcacc	30900
ggcactcgtg	ccgctggggc	ggctggatcc	gtgttctccg	gtcaagggtg	tcagtggccc	30960
gggatgggaa	gcgggctcca	ctcggcggtt	ccggtgttcg	ccgacgcgtt	tgacgaagcc	31020
tgctgcgagc	tggatgcgca	tctcggggcag	atggcccggc	tacgagatgt	gttgtccggg	31080
tcggatacgc	aacttctgga	ccagaccttg	tgggcgcagc	cgggcctggt	cgcgttgcaa	31140
gtcggactct	gggagttggt	gggttcgtgg	ggtgtccggc	ccgctgtggt	gctgggccac	31200
tcggtcgggt	agctggcggc	ggcgttcgcg	gctggagtgt	tgctggttgc	ggatgcggct	31260
cggctggtgg	cgggcctgtc	ccggttgatg	caagccctgc	caactggcgg	tgccatgctc	31320
gctgcggctg	ctggagagga	gcagctgcgc	ccgttgctgg	ccgactgcgg	tgatcgtgtg	31380
gggatcgcgc	cggccaacgc	tcccgggtcg	gtggtgctct	ccggtgatcg	ggatgtgctc	31440
gatgacattg	ccggtcggct	ggacgggcaa	gggatccggt	ccaggtggtt	gcgggtttcg	31500
catgcgtttc	attcgcacgc	gatggatccg	atgctggcgg	agttcacccg	aatcggcccg	31560
agcgtggact	accggtcgtc	agggtcgccg	atcgtgtcga	cgttgacggg	tgagctcgat	31620
gaggtcggca	tgccggctac	gccggagtat	tgggtgcgcc	agggtgcgag	accggtccgc	31680
ttcgcgcgac	gtgttgctgc	gctcgcggct	cacggtgtga	gcaccgtcgt	cgaggtcggg	31740
ccggatgggg	tggttgctgg	gctggtgcag	gagtgcgcgg	ccggatccga	tcagggcgga	31800
cgggtggccg	cggttccgct	catgcgcagc	aatcgcgacg	aggcgcacac	ggtgacaacg	31860
gcattggcgc	agatccatgt	gcgtggtgct	gaggtggact	ggcgggtcgt	tttcgccggg	31920
accggggcaa	agcaggtcga	gctgcccacg	tatgccttcc	aacgacagcg	gtactggctt	31980
gactcaccat	ccgaaccggg	cgggcaatcc	gccgatcccg	cgcgccagtc	gggcttctgg	32040
gaactcgtcg	agcaggaaga	tgacagcgcg	ctcagcgccg	ctctgcacat	taccggcgat	32100
cacgacgtgc	aggcgtccct	ggaatcgggt	gttcgggtcc	tctcctcctg	gcacgcggcg	32160
atccgcaacg	aatccctggt	gcaccagtgg	cggatccgga	tttcctggca	tgagcgggca	32220
gatttgccag	accctcgttt	gtcggggaca	tggctcgtcg	tcgtgccgga	ggggtggtcg	32280
gcgagtcggc	aagttctgcg	tttcaacgag	atgttcgagg	aacgggggtg	ccgggcagtt	32340
ctgttcgagc	tcgccgggca	cgaagaggaa	gccctggcgc	aacgattccg	ctcgttgctt	32400
gttgctgcag	ggggaataag	cggcgtgttg	tccttgctgg	cgttggtatg	atcgcctgcc	32460
tcgccgaacg	ctgctttgcc	gaatggcgcg	ctgaactcgt	tggtagtctg	gcgagctctg	32520
cgggcgcgcg	atgtgtcggc	gccattgtgg	ttggcgacgt	gtggtggtgt	cgcggtcggg	32580
gatgtgcccg	tgaaccgggg	gcaggcgtcg	gtgtggggac	tgggtcgctg	cggtggtctg	32640
gagcatcccg	cctggtgggg	tggcctgggt	gacgtgccgt	gcttgctcga	tgaggacgct	32700
cgagaacgct	tgctggtcgt	gttggcagggt	cttggcgagg	acgagatcgc	ggtacgtccc	32760
ggtggtgtgt	tcgtgcggcg	gttggaaacg	gctggtgcgg	cgtcgggtgc	cgggtcgggtg	32820
tggcgtccct	gggggacggg	gttgggtgacg	ggtggtacgg	gcgggtttgg	ggcgcatggt	32880
gcccgggtgg	tggcgggtgc	cggggctgag	catgtggtgt	tgaccagccg	tcgaggcgcg	32940
gcggctcccg	gcgctggaga	tttgcggggc	gagctggagg	cgtcgggcgc	tcgggtttcg	33000

atcacggcct gcgacgtgac cgatcgtgac gctttggccg aagtgtggc gaccattccg 33060  
 gatgattgcc cgctgaccgc ggtgatgcat gcggcggggg tcgttgaagt cggcgacgtg 33120  
 gcgtcgatgt gtttgaccga ctctgttggg gtgctgtcgg cgaaggcagg tggcgccggc 33180  
 aatctcgatg agttgctcgc cgatgtcgag ctggatgcct tcgtgctgtt ctcatccgtc 33240  
 tcgggtgtgt ggggtgctgg cgggcagggc gcttatgcgg cggcgaatgc ctacttggat 33300  
 gcgttggcgc agcagcgtcg ggcaaggggg ttggtgggga ctgcggttgc gtggggcccg 33360  
 tgggcccgtg acggaatggc cgcaggtgaa ggcggtgcac agctgcgccg ggccggcctg 33420  
 gtgccaatgg ctgcggatcg ggcgttgcgt gcacttcagg gcgcattgga tcgtgacgag 33480  
 acatccctgg tcgtggccga tatggcgtgg gagaggttcg ccccggtgtt cgccatgtcc 33540  
 cgctggcgtc cgctgctcga cgagctgccc gaagcacagc aggcgttggc ggatgcggag 33600  
 aacaccactg atgctgcgga ctcgccgtc ccgctaccgc ggctcgccgg catggcagcc 33660  
 gccgaacgcc gccgcgcgat gctggacctg gtgctggcgg aggcctcgat tgtgttggga 33720  
 cacaacgggt ctgaccagc ttggtccgac cgggcttcc aggagctcgg atttgattcg 33780  
 ctgatggccg tcgaactgcg caacaggttg ggcgaggcaa caggattgag tctgccggcc 33840  
 acgttgatct tcgattatcc gagcccatcc gcgctggctg agcagctggt cggcgagctg 33900  
 gtgggagcgc agcccgcgac caccgtcgtg gccggggccg atccagtggg tgatccggtt 33960  
 gtcgtggtcg cgatgggatg ccggtatccg ggcgacgtct gctcgcccga ggagctgtgg 34020  
 cagctggttt ctgcgggacg tgatgcggta tcgacgttcc ccgtcgatcg gggttgggac 34080  
 tgcaacacgt tgttcgaccc ggatccggat cgggcaggca gtacctatgt gcgagaaggt 34140  
 gccttctga ccggtgctga tcggttcgac gccgggttct tcggcatcag ccctcgcgag 34200  
 gcgcgcgcaa tggatccgca gcagaggttg ttgctcgaag tggcgtggga ggttttcgaa 34260  
 cgagcaggaa tcgctccgct gtcgttcgcg ggtagcagga ccggtgtgtt cgcggggacc 34320  
 aatgggcagg accacggtgc gaaagtggct gccgcgccgg aggcggccgg tcacctctg 34380  
 accggaacg ccgcgagtgt cctggccggc cggctttcct acacgttcgg ccttgagggg 34440  
 cctgcggtgg cgggtgatac cgcgtgttcg tcgtcgttgg tggcgttgca tttggcgtgc 34500  
 cagtcgctgc gttcgggtga gtgtgatatg gcgttggcag gtggtgtgac ggtgatgtcg 34560  
 acacccttg ctttctcga gttctctcgt cagcgcggtt tggcgccaga tggctcgtgc 34620  
 aagtcgtttg cggccgctgc ggatggcacc ggggtgggtg aggggtgcgg cctggtgttg 34680  
 ctggagcggg tgtcggatgc tcgtcggaat ggtcaccggg tgttggccgt ggttcgcggg 34740  
 tctgcggtga atcaggatgg tgcgtcgaat ggcctgactg cgcgaatgg tccgtcgag 34800  
 cagcgggtga ttcggcaggc cctcgcgaat gcggggctgt cggcgtccga tgtggatgtc 34860  
 gtggaggcgc acgggaccgg taccgggctc ggggatccga tcgaggcgca ggcgctgac 34920  
 gcgacatatg ggcaggagcg ggatcctgag cgggccctgt ggctggggtc gatcaagtcc 34980  
 aacatcggcc acacgcaggc ggcggccggt gtggcggggg tcatcaagat ggtgcaggcc 35040  
 atgcggcacg gggagtgtcc tgcgacgttg cacgtggaca agccactcc acaggtggac 35100  
 tggctcgcgg gggcgttcg gtcctcacc gggaacacgc cctggcccga gagcggccgt 35160  
 cctcgtcgag cgggggtgtc gtcgttcggg atcagcgga ccaacgcaca cctcatctc 35220  
 gaacaaccac cgtcggaacc agcggagatc gaccaatcgg atcggcgggt cactgcgcat 35280  
 ccagcgggtg tcccggtgat gttgtcggct aggagtctcg cagcgtgca ggcccaagcg 35340  
 gctgcgtgc aggcggggt ggaccggggt cctggcgctt ctccgctgga tttggggtat 35400  
 tcaactcgca ccaactgttc tgtgctggac gaacgcgcgg tcgtgtgggg tgccgatcgg 35460  


---

 gaggcactgc tgtccaggt ggcagcgtc gccgatggc ggaaggcgcc gggggtgata 35520  
 acgggctctg cgaattccgg tggccgcac ggattcgttt ttccgggtca gggcagtcag 35580  
 tggctgggga tgggaaaggc gttgtgcgcg gctttccgg cgttcgcgga cgccttcgag 35640  
 gaagcctgcg acgcgctaag cgcacacctg ggcgcggacg ttccgggtgt gctgttcggt 35700  
 gctgatgagc agatgctcga ccggacgtg tgggcgcagt cggggatctt cgcggttcaa 35760  
 gtgcgcctcc tgggattgct gaggtcgtgg ggcgtgcggc cggccgcggg gctggggcac 35820  
 tcggtcggcg agttggctgc ggcgcacgc gctggtgtgt tgccttgcc ggacgctgca 35880

cggttggttg	cggttcgggc	ccacctgatg	caggcattgc	ccaccggcgg	cgcaatgctc	35940
gcggtcgcca	ccagcgaggc	ggcggtcgga	ccgctgcttt	ccggggtgtg	cgatcgggtc	36000
agcatcgctg	cgatcaacgg	ccccgagtcg	gtagtgtctt	ccggcgaccg	cgatgtgctc	36060
gtggagctcg	caggcgaatt	cgatgcccga	gggcttagga	ccaaatggtt	gcgggtctcc	36120
catgctttcc	actcgaccg	gatggaaccg	attctggacg	agtacgcgga	aaccgccagg	36180
tgcgtcgagt	tcggtgaacc	ggtggtgccg	atcgtctccg	ccgcgaccgg	tgcgtggac	36240
accaccggac	tgatgtgcgc	ggccgactac	tggacgcgcc	aagtgcgtga	tcctgtccgc	36300
ttcggagacg	gtgtccgggc	gctcgtcggc	caaggcgtgg	acacgatcgt	cgagttcggc	36360
ccggacgggg	cgttgtcggc	cctggtcgag	cagtgccttg	ccgggtccga	ccaggctggg	36420
aggggtggcg	cgatcccgt	gatgcgcagg	gaccgcgatg	aggtcgagac	cgcggtggcg	36480
gccctggcgc	acgtgcacgt	ccgcggtggt	gcggtggact	ggtcggcttg	cttcgccggc	36540
accggcgccc	gcaccgtcga	gttgcccacc	tacgccttcc	aacgccagcg	gtactggctg	36600
gccgggcaag	cggacgggcg	cggcggcgat	gtggttgccg	acccggtcga	cgcgcgcttc	36660
tgggagttgg	tcgagcgcgc	cgatccggaa	ccgttggtgg	atgaactctg	catcgaccgg	36720
gaccagccct	tccgggaggt	gctgcccgtt	ctggcttcc	ggcgcgagaa	acaacgccag	36780
gaggccctcg	cggattcctg	gcgctaccag	gtgcgctgga	ggtccgctga	ggtgccgtcc	36840
gcagccgccc	tccggggcgt	gtggctggtg	gtgcttcag	ctgacgtgcc	ccgagatcaa	36900
ccggcggtcg	tcacgcacgc	gctgatcgcg	cgcggcgcgc	aggtcgcggt	cctggaattg	36960
accgagcagg	acctccaacg	cagtgcgctt	gtggacaagg	tgcgcgcgt	cattgcggac	37020
cgcaccgagg	tgacgggtgt	gttgtctctg	ttggcgatgg	acggcatgcc	ctgcgcggcg	37080
catccgcacc	tgtcccgtgg	tgtcgcgcgt	accgtgatcc	tgacgcaggt	gttgggcat	37140
gcgggtggtt	ccgcccgcgt	gtggctggcc	acgaccggtg	gcgtcgaggc	cgggaccgag	37200
gacggtccgg	ccgatccgga	ccacggcttg	atctgggggc	tcggcagggg	cgtcggcctt	37260
gaacatccgc	agtgtggggg	tggcctgatc	gaccttcgg	agacactgga	cgagacgtcc	37320
cggaacgggt	tgggtggcgc	actcgccggg	acggcggcgc	aagatcagct	cgccgtgcgt	37380
tcacccgggt	tgttcgttcg	cagagtgttg	cgcgcacgcg	ggaacccccg	gtcagagaca	37440
tggcgtagcc	ggggaacggt	cctcatcacg	ggcggaacag	gcgcgctcgg	tgccgaggtc	37500
gcacgatggc	tggcccggcg	gggagctgag	cacctggtgt	tgatcagtcg	ccgcggcccc	37560
gaagctcccg	gcgcagcgga	cctagggggc	gagctgactg	aactcggcgt	gaaagtcaca	37620
gtcttgccct	gcgatgtgac	ggaccgcgac	gagctggcgg	cggtgctggc	ggccgttccc	37680
acggagtatc	cgctgtcggc	ggtcgtgcac	accgccggcg	tcgggacgcc	tcggaacctg	37740
gccgagacga	ccttggcgca	gttcgccgac	gtgttgctcg	ccaaggctcg	cggcgcggcg	37800
aacctggacc	ggctgcttg	cgggcaaccg	ttggacgcct	tcgtgctggt	ctcctcgatc	37860
tcgggagttt	ggggagccgg	cggccaagga	gcctattcgg	ccgccaatgc	gtatctcgat	37920
gcccttgccg	agcgccgacg	ggcttgccgg	cggccggcga	cgtgcacgcg	ctggggtccg	37980
tgggcgggtg	cgggcatggc	cgttcaggaa	ggtaacgagg	cgcatctccg	ccgaaggggc	38040
ctggtaccga	tggaaccgca	gtcggccctc	ttcgcgctgc	aacaggccct	gtcccaacga	38100
gaaaccgcca	tcaccgtcgc	agatgtggac	tgggagcgat	tcgccgcctc	tttcaccgcg	38160
gcccgcgccg	gaccactggt	ggaagagatc	gtggatctac	ggcccgacac	cgagaccgag	38220
gagaagcacg	gtgccggcga	gctggggcag	cagctggccg	cactgccgcc	cgctgagcgc	38280
ggacacctgc	tgctggaggt	ggtgctggcg	gaaaccgcca	gcaccctggg	gcacgattcg	38340
<del>gaggaggetg</del>	<del>tgeaaeeega</del>	<del>teggaeette</del>	<del>gecgaaetgg</del>	<del>gettegatte</del>	<del>getgaccgcg</del>	<del>38400</del>
gtagagctgc	gcaacaggtt	gaacgcggtg	accgggcttc	gcctgccgcc	gacgctggtt	38460
ttcgaccacc	cgacgccgct	ggcgttgctc	gaacagttgg	ttccggccct	ggtcgcggag	38520
ccggacaacg	gcacgaatc	gctgctcgcc	gagctcgaca	ggctggatac	cacgttgggc	38580
caagggcctt	cgatcccact	ggaagaccag	gccaaggtgg	cggagcgctt	gcacgcactc	38640
ctcgccaagt	gggacggggc	gcgtgacggc	acggccagag	cgacgtcacc	ccaatcgctg	38700
acggcggcca	cggacgacga	aatcttcgac	ctcatcgacc	ggaagtccg	gcgctgaccg	38760

ccctttctct gcctcag ccttgattac tggaacggtg tattttctg gccaatgaag 38820  
aaaagctccg cgagtacctc aagcgtgtcg tcgtcgaact ggaagaggcg cacgaacgcc 38880  
tgcacgagtt ggagcgccag gagcacgacc ccatcgcgat cgtgtcgatg ggatgtcgtt 38940  
atcccgggtg cgtctccact ccggaggagc tgtggcgact ggtcgtcgac ggaggagacg 39000  
cgatcgcgaa cttccccgaa gaccgtggct ggaatctgga cgagctgttc gatcctgac 39060  
cgggcccagc cgggacctcc tacgtccgag aggggtggtt cctgcgcggg gtcgcggact 39120  
tcgatgcggg gctcttcggg atcagtccgc gcgaggcaca ggcgatggac ccgcaacagc 39180  
ggttgctgct ggagatctcg tgggaggtgt tcgagcgcg cggcattgac ccgttttctt 39240  
tgcggggtac caagaccggt gtgttcgagg gctgatcta ccacgactac gcgtcgcgg 39300  
ttcgcaagac ccccgaggag ttcgaggggt acttcgccac cggcaacgcg ggcagcgtcg 39360  
catccggccg ggtggcttac accttcgggt tagagggccc ggcggtcacc gtggacaccg 39420  
cctgctcgtc gtccctgggt gcgctgcacc tggcctgcca gtccctgcgg ctgggcgaat 39480  
gcgacctggc cctggccggt ggcatttcgg tgatggccac gccgggagcc ttgctcgagt 39540  
tcagccggca acgcgcactc gcctcggatg gccggtgcaa gcccttcgag gatgccgccg 39600  
acggcaccgg ctggggcgag ggcgcgggaa tgctgctgct ggaacggctg tcggacgcac 39660  
gacgaaacgg ccaccgggtg ctggcgcgcg tggtcggttc cgcgatcaac caggacggga 39720  
cgtccaacgg cctgaccgag cccagcggtc ccgcacagca gcgagtgatc cgccaagccc 39780  
tggcgaacgc cgggttgctg cccgccgagg tcgatgtggt cgaggcgcac ggcacgggca 39840  
cggccttggg cgacccgatc gaggcgcagg ccctgatcgc cacctacggg gcgaaccggt 39900  
cggcggatca tccgctgctg ctgggttccc tcaagtcgaa catcggccac acccaggctg 39960  
ccgcgggtgt ggccgggggtg atcaagtccg tcctggccat caggcaccgg gagatgcccc 40020  
gcagccctga catcgaccag ccatcgacgc acgtggactg gtcggcgggc gcggtgcggc 40080  
tgctcacgga cagcgttgac tggccggatc tcggcaggcc gcgcgcgagca ggggtgtcct 40140  
cgttcggcat gagcgggtacc aacgcacacc tgatcgctga ggaagtatcc gacgagcccg 40200  
tctcgggcag taccgagccg accggggcat ttccctggcc gctgtccggc aagacggaga 40260  
cggcattgag cgagcaggct gccgagttgc tctccgtagt gaccgagcac ccggagcccg 40320  
gactggggga cgtcgggtac tcgttgcca ccggtcgcgc tcgatggag caccgggctg 40380  
tcgtggttgc cgacgatcg gactctttcg tcgccggact gacggcgttg gctgcgggag 40440  
ttccggcagc caacgtggtg cagggcgcgg ccgactgcaa gggaaaggtc gcgttcgtgt 40500  
tccccggcca gggctcgcat tggcagggga tggcgagggg actgtccgaa tcctgcggcg 40560  
tgttccggcg gaagctggcg gaatgcgcgg cggctacggc cccttacgtg gactggtcgc 40620  
tgctcggcgt ccttcgcggt gatcccgatg caccgcgct ggatcgcgac gacgtgattc 40680  
agctcgcgct gttcgccatg atggtgtcgc tggccgaact gtggcgttcg tgccgagtg 40740  
agcccgccgc ggtggtcggc cattcccagg gcgagatcgc cgccgcccac gtggcaggcg 40800  
ctttgtcctt gactgatgag gtgcgcacat tcgctgcccg ctgcgatgag gtgtcggcgc 40860  
tgaccgggaa gggaggcatg ctgcgattg ccttgccgga aagcgcgggt gtgaagcgaa 40920  
tcgcaggcct gccggagctg accgttgccg cggtcacacg acccggtcc actgtcgttt 40980  
ccggcgaaac gtcggctctg gagcgtctgc agaccgaact gaccgcggaa aacgtgcaga 41040  
ccggcggggt gggaattgat tacgcctcgc attcgccgca gatcgcgag gtccagggcc 41100  
ggcttctgga ccggctgggc gaagtcgggt ccgaacctgc tgagatcgct ttctactcga 41160  
cggtcaccgg cgagcggagc gacaccggcc gactcgacgc cgactactgg taccagaacc 41220  
ttaggeagee cgtcegette cagcagaccg tcgcccggat ggcagatcag ggctatcgg 41280  
tcttcgtcga ggtgagccc caccgcgtgc tcaccgccc aatccaggaa acgctggaag 41340  
ccgcggacgc gggcggggtg gtggtcgggt cgctgcggcg tggcgagggc ggctcccggc 41400  
gctggctgac ttcgctggcc gactgccagg tcgcgcggact gccggtgaat tgggaacagg 41460  
tattcctcaa caccggagcc cgacgcgtgc cgctgcgac ctaccggttc cagcggcagc 41520  
ggctactggt ggagtccgcc gactacgagc cgggcgatct cggttcgggt ggcttgctct 41580  
ccgcgcgagca tccctgctc ggggctgcgg tgacgctggc cgatgcgggc gggttcctgc 41640

tgaccggcaa	gctgtcggtc	aagaccacgc	cctgggtggc	cgaccacgtg	gtcggcgggg	41700
cgatcctgct	gcccggcacc	gcgttcgtgg	aaatgctgat	acgcgccgcg	gaccaggctc	41760
ggtgcgatct	gatcgaggag	ttgtccctga	cgactccgct	ggttttgccc	gcgaccggtg	41820
cggtgcaggt	gcagatcgcg	gttggcggtc	cggacgaggg	cgggcgccgc	tcgggtccgc	41880
tgcattcctg	tcgagacgac	gccgtgccgc	aggactcggt	gacctgccac	gcgaccggca	41940
cgttgacctc	cagcgatcac	caggacgcgc	gccagggccc	cgatgggatt	tggccgcccc	42000
acgatgctgt	cgcggttcgc	ctggacagct	tctacgccc	cgagctgag	cggggcttcg	42060
atttcggccc	ggcggttcag	gggttgacag	cggcttgga	gcgcggagac	gagatcttcg	42120
ccgaggtcgg	cctgcccacc	gcacaccgcg	aagacgcgcg	cagggttcgga	atccaccctg	42180
ctctgctgga	tgcggcactg	caggcgctgg	gcgcagccga	agaggatccg	gacgagggat	42240
ggctcccgtt	cgcggtggca	ggtgtgtccc	tcaaagcgac	gggcgcactt	tccttcgggg	42300
tgcacctcgt	tcggcggggc	gcgaatgcgg	tgtcggtgtt	cacgaccgac	acgactggcc	42360
aagccgtgct	ctccatcgat	tcgctggtgc	tgcgccagat	ttcggacaag	cagttggcag	42420
cggcccgtgc	gatggaacac	gagtccctgt	tccgggtcga	ctggaagcga	atctcgccc	42480
gcgctgccaa	gccggtctcc	tgggcagtga	tcggcaatga	cgaactcgcc	cgagcctgcg	42540
gctcggcact	tggcacggaa	ctccacccc	acctgaccgg	gttggctgac	ccgccccggg	42600
acgtcgtggt	ggtgccatgc	ggtgcgtctc	gccaggactt	ggacggtgct	tcgaggcac	42660
gtgccgcgac	acaacgcagt	cttgacctga	tccaggattg	gttggcgcg	gcgcgattcg	42720
ccggatctcg	cctggtggtt	gtgacgtgtg	gtgcggcgct	gacaggctcc	gccgagggtg	42780
tttccgacct	ggtgcatgct	gcgtcgtggg	gtttgttgcg	ttcggcgag	tcggagaacc	42840
cggaccgatt	cgtgttggtc	gatgtggacg	gaaccgcgcga	atcatggcgt	gcgctcgcg	42900
cggccgtgcg	ttccggagaa	ccgcagctgg	cgttgccgcg	cggatgaagtc	cgggtgcctc	42960
gcctggcgcg	atgtgttgcc	gccgaggaca	gccggatccc	agtcccgggt	gcggatggga	43020
cgggtgtgat	ttccggcggt	acgggcctgc	tggcggggtt	ggttgcccgg	catttggtgg	43080
cggagcgcg	tgtccgcgc	ctggtgctcg	cggggcgagc	cggctggagc	gccccgggg	43140
tcaccgacct	ggtggatgag	ttggtgggcc	tgggagctgc	ggtcgagggtg	gcgagctgcg	43200
atgtcgggga	tcgggcccag	ttggaccggc	tgtgacgac	gatctcggca	gagttcccgc	43260
tgcgcggagt	ggtgcatgcg	gccggggcac	ttgccgacgg	ggtcgtcgag	tcgctgacac	43320
cagagcacgt	ggcaaagggtg	ttcggccccga	aggccgcgcg	tgcgtggcac	ctgcacgagt	43380
tgactcttga	tctggatctc	tcgttcttcg	tgtctctctc	ctcgttctcc	ggcgtggcg	43440
gggctgcggg	tcagggaaac	tacgcggcgg	cgaacgcgtt	cctggacggc	ctggctcagc	43500
accggcgga	ggcggggctg	cctgcggtgt	cgctggcttg	gggcttggtg	gagcagccca	43560
gcgggatgac	cggagcgctc	gatgcggcgg	gccgtagccg	cattgcgcgc	accaatccgc	43620
cgatgtccgc	gccggacggg	ttgcggctgt	tcgagatggc	gtttcgcgtt	ccgggcgaat	43680
cgttcttggt	tcgggtccac	gtcgacctga	acgccctgcg	cgtgatgcg	gccgacggcg	43740
gtgtgcctgc	gttgttgccg	gacctggtgc	cagcgcgcgt	gcggcgagc	gcggtcaacg	43800
agtcggcgga	cgtcaacggt	ctggttggtc	ggctgcggag	gctgccggac	ctggatcagg	43860
aaaccagct	gttgggtttg	gtgcgcgagc	atgtttcggc	ggtgctgggg	cattcgggtg	43920
cggtcgaggt	cggggccgat	cgtgctttcc	gggatttggg	ttttgattcg	ttgtccggtg	43980
tggagtctcg	gaaccggctt	ggcggggtgc	tggcggttcg	gttgccggct	actgcggtgt	44040
tcgactatcc	gacaccgcgg	gcgttggttc	ggttcttgct	cgacaaactg	attggtggcg	44100
tgagggctcc	gactcccga	ccggcggtcg	tggcgcggtt	gactgctgac	gateccgttg	44160
tgatcgtggg	gatgggctgt	cgttatccgg	gtggggtgtc	ctcgccggag	gagctttggc	44220
gtttggtggc	cgggggcttg	gatgcgggtg	cggagttccc	ggacgatcgt	ggctgggac	44280
aggcggggtt	gttcgatccg	gatcccgatc	gtcttgggac	ctcgtatgtg	tgtgagggtg	44340
gcttccctgc	agatgcggca	gagttcgatg	ccggtttctt	cgggatttcc	ccgcgtgagg	44400
cgttggcgat	ggatccgcag	cagcggttgc	tgttggaagt	cgttgggaa	accgtggagc	44460
gggcggggat	tgatccgctt	tcgttgccgg	ggagccggac	cggcggttgc	gcggggctga	44520

tgcaccaoga ctacggcgcg cggttcatca cgagggcgcc ggaggggttc gaggggtatc 44580  
 taggtaatgg cagcgcggga ggcgtgtttt cgggtcgggt tgcgtattcg tttgggttcg 44640  
 agggctcctgc ggtgacggtg gatacggcgt gttcgtcgtc gttgggtggc ctgcacctgg 44700  
 cgggtcaagc actgcggtct ggtgagtgtg atctggctct tgcgggtggt gtgacggtga 44760  
 tggccacgcc ggggatgttc gtggagtttt cgcgtcaacg gggcttggcg gcggatgggc 44820  
 ggtgcaagtc gtttgcggcg gctgcggatg gcaccgggtg gggagaaggc gcgggcttg 44880  
 tgttggttga gcggtctgtc gatgccggc gcaacgggca cgcgggtctg gcggtcgtgc 44940  
 ggggtagcgc ggtgaatcag gatggtgcgt cgaatggttt gacggcgccg aatgggcct 45000  
 cgcagcagcg ggtgatcag caggcgttg cagtgctgg tttgtcgggt tctgatgtgg 45060  
 acgccgtgga ggcgcattgg actggaacca ggcttggtga tccgattgag gcgcaggctc 45120  
 tgattgccac ttacgggcag gggcgggata gcgatcgcc gttgtgggtg gggtcgggtga 45180  
 agtcgaatat tggtcatacg caggcggcg cgggtgtcgc tgggtgtgac aagatgggtga 45240  
 tggcgatgcg gcacgggcag ctgcccgca cgttgcatgt ggatgaacct acgtcgggaag 45300  
 tggattggtc ggcgggggat gtccagctcc tcacggagaa cacccttg cccggcaaca 45360  
 gccatcctcg gcgggtgggc gtgtcgtcgt tcgggatcag cggcaccaac gcacacgtca 45420  
 tctcgaaca agcctcgaaa acaccagacg agactgcgga caagagcggc cccgattcgg 45480  
 aatcgaccgt ggacctcca gcggtcccgt tgatcgtgtc ggggagaaca ccggcagcgc-45540  
 tcagcgtca ggcgagcgca ttgttgtct atttgggtga gcgtggcgat atttccacgc 45600  
 tggatgcggc gttttcgttg gcttctccc gggcgcggtt ggaggagcgg gcggtggtgc 45660  
 tgggagcggga ccgcgaaacg ttgttgtcc ggttggaagc gctggcttc ggtcgcgagg 45720  
 cttctggggg ggtgtcggga tcccgggtct ctggcgggtt tgggttcgtg ttcgccggtc 45780  
 agggcggaca gtggttggg atgggcccgg ggcctactc ggtttttccg gtgttcgctg 45840  
 acgcgttga cgaagcatgt gccggactgg acgcgcactt ggggcaggac gtgggggtcc 45900  
 gggatgtggt gtttggttc gacgggtcct tgggtgatcg gacgctgtgg gccagtcgg 45960  
 gtttgttcgc gttgcaggtt ggtttgctga gcctgctggg ttcgtggggg gtccggccgg 46020  
 gtgtggtgct gggccattcg gtcggcgagt tcgcggcggc ggttgcggcg ggagtgttgt 46080  
 cgttgccgga tgcggctcgg atggtggcgg gtcgtgccg gttgatgcag gcgttgctt 46140  
 ctggcgggtg catgttggtg gtggctgctg gtgaggagca gctgcggccg ttgttggtcc 46200  
 atcgggttga tgggtgcgggt atcggcggcg tcaacgctcc tgagtcggtg gtgctctccg 46260  
 gcgatcggga ggtgcttgac gacatcgccg gcgcgctgga tgggcaaggg attcgggtgg 46320  
 ggcgggttgc ggtttcgcgt gcgtttcatt cgtatcggtt ggaccgatg ttgcaggagt 46380  
 tcgccgaat cgcacgcagc gtggactacc ggcgtggcga cctaccggtc gtgtcgacgt 46440  
 tgacgggtga gctcgacacc gcaggtgtga tgggtacgcc ggagtattgg gtgcgtcagg 46500  
 ttcgagagcc cgtccgcttc gccgacggcg tcgggtgct cgcgcagcaa ggggtcgcca 46560  
 cgatcttcga actcggccct gatgcgacgc tgcggccct gattcccgat tgtcattcgt 46620  
 gggctgatca ggccatgccg attccgatgc tgcgtaaaga ccgtacggaa accgaaactg 46680  
 tggtcgccg ggtggcgcg gcgcacacgc gtggtgttcc ggtcgaatgg tcggcgatt 46740  
 tcgccggcac cggggcacgg cgggtcgagt tgccgacgta tgccttcag cggcagcgg 46800  
 actggctgga aacatcggt tacggcgatg tgacgggtat cggcctggct gcggcgagc 46860  
 atccgttgct gggggccgtg gttgcgctgg ccgatggtga tgggatggtg ctgaccggcc 46920  
 ggttgctcgtt ggggacgat ccgtggctgg cccagcatcg cgtgctgggc gaggtcgtcg 46980  


---

 tccccggaac cgcctcctg gagatggccc tgcacgcagg ggcgcgtctc ggtgtgacc 47040  
 ggggtggaaga gctcaccctg gaaacaccgc tgggtggtccc cgaacgcgcg gcgggtgccg 47100  
 gtagtcgtgg ccctgcggga gggaccacag tttcaattga aactgcggaa gaacgtgtgc 47160  
 ggacgaacga cgcctacgaa atccagctgc tgggtgaacgc acccgacgaa ggcggtcggc 47220  
 gaagggtgtc gctgtattcc cgcggggcgg gtgggtcgag aggtgggggt tggacgcgc 47280  
 acgccaccgg cgaactcgtc gtcggcacca ccggtggtag ggcggttcct gattggtcgg 47340  
 ctgaggggtg cagatcgatt gctctcgatg agttctacgt cgtctggcc ggaaacgggt 47400



tcgagtacgg gccgttggtc caggggcttc aggcggcatg gcgtcgtggt gacgaggttc 47460  
 tcgccgaat cgcggcgccg gccgagccg atgcgatggc gtcgggatac ctgctcgacc 47520  
 cagcgttgct ggatgcccg ctcgagcggt ccgcgctcgg cgaccgcccg gagcaaggcg 47580  
 gcgcgtggct gccgttctca ttcaccggcg tcgaactttc cgctccggca gggacgatca 47640  
 gcaggggtcg gctggagacc aggcgacccg acgcgatata ggtggccgtg atggatgaga 47700  
 gtgggcggtt gctcgccctc atcgattctc tcaggctacg aagcgtgtcg tcgggacagc 47760  
 tggcgaatcg ggacgctgtc cgcgacgcgc tgttcgaggt gacctgggag ccggtggcga 47820  
 cgcagtcgac ggaaccgggt cgctggggcc tgcctgggtga tactgcctgc ggtaaagacg 47880  
 atctcatcaa actcgcaacg gattccgccg accgctgcgc ggatctggcg gcgctagccg 47940  
 agaaacttga ttccagcgcg ctggttcctg atgtcgtggt ctactgcgcc ggagaacagg 48000  
 cggatcccgg caccggcgca gccgcacttg cggagaccca gcagacgttg gctctgctcc 48060  
 aagcgtggtt ggctgagccg cggttggccg aggcacgtct ggtggtggtg acgtgtgcag 48120  
 cggtgacgac ggctccgagt gacggtgcat cagagctggc acatgcgccg ttgtgggggt 48180  
 tgttgctgct gcgcgaggtg gagaacccgg ggcagtttgt gctggcggac gtcgacggaa 48240  
 ccgccgaatc gtggcgtgct ttgccgagtg cgttgggctc gatggaaccg cagttggccc 48300  
 tgcggaaggg cgcggtgcga gcgccccgct tggcttcggt cgccgggcag atcgacgtgc 48360  
 ccgcggttgt ggcgatccc gaccgaaccg tgcgtgattc gggcggcacg ggcctgttgg 48420  
 gggcgcggtt tgcccgcac ctggtgaccg aacgcggtgt ccgccgattg gtgttgacgg 48480  
 gccgtcgtgg ctgggatgct cctggaatca ccgagttggt gggtagctg aacggcctcg 48540  
 gtgccgtggt cgacgtgggt gcgtgcgacg tcgcggatcg tgcgtatctg gactcgttgc 48600  
 tggcggcggt ccgcggcgaa tttccgttgt gcggcggtgt gcatgcccg gggcgctgg 48660  
 ccgacggggt gatcgagtcg ttgtcaccgg acgacgtggg agcgggtgtc ggcccgaagg 48720  
 cggcgggggc gtggaatctg cagagctga ctcgtgatac ggacctgtcg ttcttcgcgt 48780  
 tgttctctc gctttccggt gttgccggcg ctctggtca gggcaattat gcggcgcgca 48840  
 acgcgttcct ggacgcattg gcgcattacc ggcggtcaca gggactgcct gcggtgtcgc 48900  
 tggcctgggg cctgtgggag cagccgagcg ggatgacgga gacgctcagc gaggtcgacc 48960  
 ggagcaggat cgcgcgcgcc aaccgcgctt tgtccaccaa ggagggattg cgctgttcg 49020  
 atgccgggct ggcgctggac cgggcagcgg tagttccggc gaagttggac aggactttcc 49080  
 tggccgagca ggcgcggtcg ggctcgtcgc ccgcattgtt gacggcactg gtacccccca 49140  
 tccgtcgtaa taggcgggct agcggaaacc agctcgcgga cgagggcacc ctgctcgggg 49200  
 tggtagggga gcatgcccg gccgtgctgg ggtattcgag cgcggctgac gtcggggtcg 49260  
 agcgcgttt ccgggatctg ggttttgatt cgttgtctgg tgtggagttg cggaaccgcc 49320  
 tgccgggggt gctgggggtg cggttgcgg cgactgcggt gttcgactat ccgacgccga 49380  
 ggcgctggc ccggttcctg caccaggaac tggcagacga gatcgctacg accccagcgc 49440  
 cggtagcgac gaccaggga ccggtcgccg aagacgatct cgtcgcgata gtcgggatgg 49500  
 gatgccgttt tcccgttcag gtgtcctcgc cggaggagct ctggcgtttg gtggccgggg 49560  
 gcgtggatgc ggtcgcggac ttcccagccg atcgcggctg ggatctggca ggcttgttcg 49620  
 atccggaccc ggaacgggtt gggaagacct acgtgcggga aggggccttc ctaccgacg 49680  
 ccgatcgggt cgatgcgggt ttcttcggga tttcccgcg tgaggcggtg gcgatggatc 49740  
 cgcagcaacg gctgttgctg gagctgtcct gggaggccat tgaacgggca gggatcgatc 49800  
 cgggttcgct gagggggagt cggaccggtg tgttcgcggg gctgatgtac cagactatg 49860  
 gcgcccgggt ccgcagccga gcccgggaag gtttcgaggg gtatctcggc aatggcagtg 49920  
 ctgggagtg cgcgtcgggc cggattgcgt actcgtttgg tttcgagggt cctgcggtga 49980  
 cggtaggatac tgcgtgttcg 50000

<210> 6

<211> 29736

<212> DNA

<213> *Saccharopolyspora spinosa*

<400> 6

tgcgtcggttg tggcggttgca tttggcggggt cagtcggttg gttccggcga atgcgatctc 60  
gcccttgccg gtggtgtgac ggtgatgtcg acgcccggga cgtttgtgga attctcccgt 120  
cagcggggcc tggcaccgga cgggcgggtgc aagtcggtcg cggagagcgc ggacggtacc 180  
ggttgggggtg aggggtgctgg tttggtgttg ttggagcgggt tgcggtatgc tcggcggaat 240  
gggcatcggg tgttggcggt ggttcgtggg tcggcggtga atcaggatgg tgcgtcgaat 300  
ggcttgaccg cgccgaatgg tccctcgag cagcgggtca tccagcaggc gttggcgagt 360  
gcgggtctgt cgggtgtccga tgtggatgcc gtggaggcgc atgggaccgg gaccagggtg 420  
ggtgatccga ttgaggcgca ggctctgatt gctacgtatg ggcgcgatcg tgatcccgtt 480  
cggccgttgt ggttgggggtc ggtgaagtcc aacatcggtc atacgcaggc ggcgcggggt 540  
gttgccggtg tgatcaagat ggtgatggcg atgcggcacg ggcaacttcc gcgcacgctg 600  
cacgtggatg caccctctc gcaggtgat tggtcggcgg ggagggtcca gctcctgacg 660  
gagaacacgc cctggcccga cagtggtcgc ccctgtcggg tgggggtgtc gtcgttcggg 720  
atcagcggca ccaacgcgca cgtcatcctg gaacagtcca cggggcagat ggatcaggca 780  
cggagccgg attcgagtcc tgttctggat gttccggtgg tgccgtgggt ggtgtcgggc 840  
aaaacacccg aagcgctatc cgcccaggcg gcaacgttgg cgacctattt ggacaaaaat 900  
gttgatgtct cccctctgga cgttgggatt tcgcttgccg tgaccggttc ggcgctggat 960  
gagcggggcg tgggtgctggg gtcggatcgt gacacgttgt tgtctggcct gaatgcgctg 1020  
gctgccggtc atgaggtgct tggcggtgtt acgggacctg tcgggattgg tggccggacc 1080  
gggttttgtt tcgccggtca aggcggtcag tggttgggga tgggcccggc gttgtactcg 1140  
gagtttccgg cgttcgccgg tgctttcgac gaagcatgcg ccgagctcga tgcgaacctg 1200  
gggagggaa tgggggttcg ggatgtgggt ttcggctccg acgagtcctt gctggatcgg 1260  
actttgtggg cgcagtcggg tttgttcgag ttgcaggctg gtctctggga attgttgggt 1320  
acgtgggggt ttcggcccag cgtagtgtcg gggcattcgg tcggggagct agccgcggcg 1380  
ttcgccgcag gtgtgctgtc gatggcggag gcggctcggc tggtgccggg tcgtgcgcgg 1440  
ttgatgcagg cgttgccctc tggcggtgcc atgctggcgg tgtccgcgac cgaggcccga 1500  
gtcggcccg cgtcgtatgg ggtgcgggat cgtgttggtg tcgcagcggg taacgctccg 1560  
gggtcgggtg tgccttccgg tgaccgggat gtgctcgatg gcattgccgg tcggctggac 1620  
gggcaaggta tccggtcgag gtggttgccg gtttcgcacg cgtttcattc gcacgggatg 1680  
atccgatgc tggcggagtt cgccgagctc gcacggagcg tggactaccg gtctccacgg 1740  
tgccgattg tctcgacgct gaccggaaac ctcgatgacg tggcggtgat ggctacgcg 1800  
gagtattggg tgcgccaggt gcgagagccc gtcogcttcg ccgacgggtg ccaggcgctt 1860  
gtggaccaag gcgtcgacac gattgtggaa ctcggtccgg acggggcggt gtcgagcttg 1920  
gttcaagagt gtgtggcgga gtccggcgcg gcgacgggga ttcggttggt gcggagagac 1980  
cgtgatgagg tccgaacggt gctggacgct ttggcgcgaga cccacactcg tggtgccgcg 2040  
gtggactggg ggtcattttt cgctggtacg agggcaacgc aagtcgacct tcccacgtat 2100  
gccttccaac gacagcggta ctggctggag ccatcggtt ccggtgatgt gaccggtgtt 2160  
ggcctgaccg gggcggagca tccgctgttg ggtgccgtgg tgccggtcgc gggcggcgat 2220  
~~gaggtgctgc tgaccggcag gctgtcggtg ggaagcctc cgtggctggt ggaacacgcg 2280~~  
gtgctggggc aagtcgtcgt ccccggcacc gcgttgctgg agatggcgtg gcgggcccgt 2340  
agccaggctc gttgtgaacg tgtggaggag ctacaccttg aggcaccgct ggtcctgccg 2400  
gagcggggcg ctgocggcgt gcagttggcg gtgggggctc cggtatgagg cgcccggcgc 2460  
agtttgcagc tctattcccg aggcgctgat gaagacggcg actggcgcg gattgcctcc 2520  
gggctgttgg cccaggccaa tgcggtgccg ccggcggtt cgacggcatg gcccgccgac 2580  
ggcgccgggc aggtcgatct ggcggagttc tacgagcgcc tcgccgagcg cggcttgacc 2640

tacgggtccgg	tattccaagg	gctccgcgcc	gcatggcgcc	acggcgacga	tatcttcgcc	2700
gaattggccg	ggtcaccaga	cgctcgggt	ttcggcatcc	acccggcgct	gctggacgct	2760
gcaactgcacg	cgatggcgct	tggtgcttcg	cccgaactcg	aagcggtct	gccgtttcc	2820
tggcgtggcg	cccagctgta	ccgcgctgaa	ggagcagcgc	ttcgggtacg	gctctcgccg	2880
ctgggctccg	gtgcagtctc	attgacgttg	gtggatgcca	cagggcgacg	agtcgctgcg	2940
gtggaatcgc	tttcgacgcg	accggtctcc	accgaccaga	tcgggtgccg	tcgcggcgat	3000
caagagcggc	tgctgcacgt	cgagtgggta	aggctcggtg	aatctgcggg	gatgtctctg	3060
acctcctgcg	cggtggctcg	tttgggcgaa	ccggagtggc	acgctgcgct	gaagaccact	3120
gggtgtccaa	tcgagtccca	tgcggaacct	gcttcgttgg	ccaccgaggt	tgccaagcgg	3180
ggttcagctc	ctgggtgcgt	catcgctccg	tgcccgcgac	cccgaagcgt	gcaggagctg	3240
ccgaccgccc	cgcaagggc	gacgcaacag	gcgatggcga	tgctgcagca	atggcttgcc	3300
gatgaccggt	tcgtcagtac	gcgctgctc	ctgctgacgc	atcgggcggt	ctccgcagtt	3360
gctggagaag	acgtgctcga	cctgggtacac	gcgcccgtgt	ggggcttggg	ccgcagcgcg	3420
caagcggagc	acccggaccg	attcgctctg	atcgatatgg	acgacgagcg	agcatcgacg	3480
acggcactcg	ccgaagcgct	gactgcggga	gaagcgcagc	tcgcgggtcg	gtcgggagtt	3540
gtgctggcgc	cccgcctcgg	ccaggtgaag	gtgagtggag	gtgaagcgtt	caggtgggat	3600
gaaggcaccg	tgctggtcac	cgggcggaacc	ggcgggctcg	gggcccgtgt	cgacgcccac	3660
ctgggtcagcg	cccacggtgt	gcggcacctg	ttgctcgcaa	gtcgcggtgg	tctggcggcg	3720
cccgagcgcg	atgagctggt	ggccgagctg	gagcaggccg	gcgcgcagct	cgcggtcgtc	3780
gcgtgcgact	cggcagatcg	ggactcgctt	gcgcggctgg	tggcgtcggt	gcctgcggaa	3840
aaccggttgc	gggtgggtgt	gcacgcccgc	ggtgtgctgg	atgacggtgt	gctgatgtcg	3900
atgtcgccgg	agcgttgga	cgcggtgttg	cggcccaaag	tggatgccgc	gtggtacctg	3960
cacgagctga	ctcgggaact	cggctctgtc	gcgttcgtgt	tgttctctct	ggtcgcgggc	4020
ctgttcggcg	gtgcggggca	gagcaattac	gctgccggca	acgctttcct	ggatgccttg	4080
gcgcattgcc	ggcaggccca	ggggctgccc	gcgctgtcgc	tggcctccgg	gctgtggggc	4140
agtatcgatg	gaatggcggg	cgacctcgct	gcggcagatg	tggagcggct	gtcgcgggca	4200
ggcattggcc	cgttttcggc	accgggaggg	ctggccttgt	tcgacgctgc	cgttggtcgc	4260
gacgaaccgt	tgctggcacc	ggtgcgactg	gatgtcgaag	caactgcgtgt	gcaggcccga	4320
tccgtgcaga	cccggattcc	ggaaatgctg	catggcatgg	caatggggcc	aagccgccgc	4380
actccgttca	cttcagggt	tgagccgttg	cacgaacggc	tggccggatt	gtcggagggc	4440
gaacgtcggc	agcaagtgt	ccagcgcgtc	cgccgcgata	tcgcgggtgt	actggggcac	4500
ggcaggtcga	gcgatgtgga	catcgagaag	cctttggccg	agctgggttt	cgactcgctg	4560
gcggccatcg	aactccgcaa	ccgtctcgct	accgccaccg	gactgcgggt	tcccgcgacg	4620
ctggccttcg	accacggcac	tgccggcgca	ctgcgccagc	acgtgtgcgc	gcagctaggc	4680
accgcgaccg	cgccggcacc	gaggcgaaac	gacgacaacg	acgccacgga	gcccgtgagg	4740
tcgctcttcc	aacaggcgta	tgccgctggc	cggatacttg	acgggatgga	tttgggtgaag	4800
gtcgtgccc	agttgcgacc	ggtgttcggt	tcgcctggcg	agctggaatc	cctgccgaaa	4860
cccgtccagc	tttcccgtgg	tcccgaagag	cttgcccttg	tgtgcatgcc	ggcgctgac	4920
gggatgccgc	ccgcacagca	gtacgcgcgg	atgcgcccg	ggttcgcgca	tgtgcgggac	4980
gtttcgtgta	tcccgatgcc	tggattcatt	gcgggagaac	cgctgcgctc	cgccatcgag	5040
gtggcggttc	ggacgcaggc	ggaggcggtg	ctgcagggaat	tcgccggggg	ctcgttcgta	5100
ctgggtcggc	attcctccgg	gggctggctg	gcgcacgagg	tagccggtga	gctggagcgt	5160
cgccggggtcg	tcccggccgg	ggtcgctactg	ctggacacct	acatccccgg	tgagatcacg	5220
ccgaggttct	ccgtggcgat	ggcccaccgg	acgtatgaga	agctcgcgac	tttcacggac	5280
atgcaggatg	tcggtatcac	cgcgatgggc	gggtacttcc	ggatgttcac	cgagtggact	5340
ccgacgccga	tcggtgctcc	gacgctgttc	gtgcggaccg	aagattgcgt	cgacgaccct	5400
gaagggcggc	cgtggacaga	tgactcctgg	cggccagggt	ggactctcgc	ggatgccacg	5460
gtccaggtgc	cgggcgacca	cttctcgatg	atggacgagc	acgccgggtc	caccgcacag	5520

gcagtcgcga gttggcga caaactcaac cagcgcaccg ctccgcaacg ctgacgggcg 5580  
tccttttagg accttctggg cggcacccgg caccocggcg gtgccgcctt ccgtgggtcca 5640  
ggctcgccga tcttgacggc gcacgatgcg cggcacgcgc gctgatcgtg attccgctgc 5700  
cgctcgtggc catcggcctg gcgaatcatg tcctttcggg caacgtcaaa cgaattcgtc 5760  
cgagcccgcga ttccgaggtg aggggcaccc ttgggtggct gagccgctca aggggtgcccc 5820  
tcacctcgaa attcgtccga tttgggcggg ggacgcaacc ccggtgggcg tgggtgctct 5880  
ttcttggtga cagagcggg agaagccgt gacacacctg agaggaaaag gggagcatga 5940  
tgctcaagcg ccaccgttg acgaccgcca tcaccggcct tctgggggga gtactgctgg 6000  
tcagcggctg cggaaccgcc gccgcacttc agtcctcgcc ggcccccggg catgacgcgc 6060  
gcaatggttg tatggcctcg ggccggggcg gcggggacat cggcacgtcg aactgctcgg 6120  
aggccgattt cctcgccacc gcgacaccgg tgaaaggcga ccccggcagt ttcacgtg 6180  
cgtacgggaa ccggtcggac aagacctgca cgatcaacgg cggcgtgccg aacctcaagg 6240  
gcgtggacat gagcaactcg ccgatcgagg acctgccggt cgaggacgtg cggcttccc 6300  
acgcgccaa ggaattcacc ctccagccc gtcagagcgc gtacgccggc attggcatgg 6360  
tcctggccga cagcggcgac ccgaacgcc atgtcctcac cgggttccag tctcgtctgc 6420  
cggacatgtc cgaggcccag ccggtcaacg ttctcggcga cggcaacgtg aagtccgccg 6480  
cgaagtacct gcgagtcagc tcgctgggtg ctaccgcaga cgagctgcgc taaaacccat 6540  
gtgagtcctg cagattcgac ctgcgcgtgc ggccgctccg gcgaagcgtc cgtacgtttg 6600  
tcgttggtgac cagcgttggt cactccggg cgcagcgtg gtacatactc aggcgtctcg 6660  
ggccgctcca acggggcctg gcatccggg ccgtcagagt cggcggcgct gacgcgttct 6720  
ctgtcgggcg ttgtcacgcc gccggcctcg aaccggctcc gccccgtcgg agccgggtgt 6780  
ccagcgcggg gtggcgccg ccggagccga cggcgccac cgcctgcccg agggcctttt 6840  
tcgaaccgac gaggaccacg accttcttg cccgggtgac cgcctgtag agcaggttgc 6900  
gctgcagcat catccaggcg cttgtggtca aggggatcac cagcacggg tattcgcttc 6960  
cctgcgaacg atggatggtc accgcgtagg cgtggaccag ttctcgtagt tctgtgaagt 7020  
cgtagtcgat gtctcgtcc tcgtcgttc gcacggtcag ggtctgtgct tcgttgctga 7080  
ggcgggacac gacgccctgc gtgccgttg acacgccgtt ggccgcttg tcgtagttgt 7140  
tgcgatctg cgtgacctg tcgccgacgc ggaagatccg tccgccgaac cgcgcgtctg 7200  
gcaggccctc cctggccggg gtgatcgct cctgcaacag ctggttcagc gcgcctgcac 7260  
ctgcggggcc tcgatgcac ggggcgagga cctgcacgtc ggtgcgcggg ttgaaccgga 7320  
acttccgcgg aatccggcg gcgacgacgt cgacggtag ctccggcggtc ggttcgcttt 7380  
cctctacgtg gaacaggaag aagtcggtca gccgtgtgt cagcgatag tccccggcgt 7440  
tgattcgggt cgcgttggtc accaccccg actcggcgcc ctgccggaac acctcgttga 7500  
ccgcacgtg tggaatcggg gtgccaggg cgagcagatc gcgcagtacc tcaccggctc 7560  
cgaccgacgg gagctggtcg acgtcgccga ccagcagcag gtgcgcgccg ggccgcatcg 7620  
ccttggccag tttgttggt aacagcaggt cgagcatgga cgcctcgtcg accacgacga 7680  
ggtcggcgtc cagcgggttg tcccggtcgt aggcggcgtc cccgccggc tggagttgga 7740  
gcaggcgggt caggtcgcc gcgtcgtgtc cggtagctc ggtcagccgc ttcgcgcctc 7800  
gtcccgtcgg cgcggcgagg atcaccttg ctttttcgc ctgagctaat gcgatgatcg 7860  
accgcacggg gaagctcttg ccgcagcctg gacctccgg gagcacggcg accttctcgg 7920  
tcagggccag cttgacggcg cgtcctcgt cctcggcgag ttcggcaccg gtagcgcggc 7980  
gcaaceagtc gagggeettg tgcaatcga cgtcggcgaa gacgggcaat cggtecege 8040  
tggtgttcag cagccgggac agctggttg ccaggcgac ttcggcgcgg tggaagggca 8100  
cgaggtagat cgcgaccgtc ggcacctcgt cgtcatcggt ggggatctcc tcgaggacca 8160  
caccttctc ggtgacgagt tcggcgaggc attcgatcac cagcccggtg tcgacggcga 8220  
ggatcttcac cgcctcggcg atcagctcgt tctccggcag gtagcagttg ccgtcgcggg 8280  
tggaactcga cagcgtgaac tgaaggcccg cttttaccg ctgcggggag tcgtcgggga 8340  
ttcccaccgc tttggcgatg gtgtcggcg tcttgaaacc gattccccac acgtcgctg 8400

ccagccggtg ttgctcttcc ttgacgggtc ggatcgcgtc gtcgtggtac tgctttaga 8460  
tcttcaccgc cagcgaggtc gagacgccga cgccttgag gaagatcatc acctccttga 8520  
tcgccttctg ctctcccccac gcgtcggcga tcagcttcgt ccgcttcggg ccgagcttgg 8580  
ggacctcgat cagccgcgcg ggttcctgct cgatgacgtc gagcgcggcg acgccgaagt 8640  
ggtcgacgat cttctcggcg agtttggggc cgatgccctt gatcaggcca gacccaggt 8700  
agcggcggtat accttgacag gtcgcaggca gcacggtcgt gtagtcgtcg acgtggaact 8760  
gccgcccgtg ctggggggtg gacccccacc ggccgcgcgt gcgcaacgcc tcgccgggct 8820  
gcgcgcccag cagcgcgccg acgaccgtca ccaggtcacc gcccgggccc gtgtcgatcc 8880  
gcgcgacggt gtagccgctc tcctcgttgg cgaacgtgat ccgctccagc gtgccctcca 8940  
gcaccgcagt ccacgtggcc gactcccgtc ctttttccac cgacaacacg tatcacgaac 9000  
ggctgtcaag caaacggcg gtcaccacat gcagcggcat ctcccgaacg cctcgggctc 9060  
cggcgtcagc ggggtgggct tcgcgatgcc ttggtgcggc cgggtggagt tgtagatttt 9120  
ttcgctcctg cgcagggcct ggagtaggtg ccgctggctc cagatcaggg gaaagcgccg 9180  
ggataaaccc gcttgacgga ggagatggaa gagctctacg tcgaatggcc agcgacccac 9240  
gacgacccc agtcatgctg cgacgacctg cgaggccgtg gcgaagcgtt gacagggggc 9300  
acgtgggggc gtgcgatcag gtcgcggcgc cggctcagcg ccgtgtccgg acgcaccagc 9360  
agccgcgacg gactggttcc gtcctcagag cgcggtgtcc tcgatcgtct ttttcgagct- 9420  
cgatgccggc cgcgttcgcg aggtggccag tccagtcggt cggcagtcgt ggtcgcagtc 9480  
tgccctttga acatctggtc aattcagctt tgaatgatcg ttcggcgagt attgtcctgc 9540  
ccgccaattg tttgtcgcgc gtgcgcctcg ggacaggagc gcggcgcttcg gggacggcgt 9600  
gctggcggaag gtgtcgattg gtcgggacgg ccggaccagg atcgccgacg tcgtgatcgc 9660  
cacgagcgtc ggcgatgccc ggtactatct ggagcaccag cgagtggaca acgacttcca 9720  
cgggcgcggt gcgctcctga tcatgaatca agaactgaga ccgtcagggt gaacgtgccg 9780  
tcgtggggca ctgtgaggat gtgcataca cggccgggggt aggtgatcga ggaggccctg 9840  
atcgcgggct gccgcggctg tgcccagtcg aatcgagca ccacctggcc gcgggtgagc 9900  
gtcagccccct gcgcgtcgag tgcgactttg tcggagtccc tcaccaccaa cgggatctgc 9960  
tcggttgccg gtgcgttcgc ctgcaccgac ctggtgatcc ggtcgggccc caggtcaca 10020  
tcggtgacga ccgagccatt cgcggtgcgg tagcggaacc ggtatggctc gccgccagc 10080  
aactccgcct gctgcggctc gttcgcgtcc ggcaacctgg ccgggaagac cgtcgcccag 10140  
caggcggtgt cgttggtgtt catcgactgg acgaccgtgc ccgcgaccgg gtgccacagg 10200  
aacgtcagac cgggtgcggc cagcgaggtc gcgcgcctcg tgccgaagta gccgccagc 10260  
tagaactgcg gcctgcgtag gaacaggaag tcttgccctt ggtcgtgcg gatctcgggtg 10320  
aagtgcgcct gttggtatgg aagtttcgcg atcgcgggcc tgcgctgggc gcgggtcgga 10380  
tagcgctcgc cgtagatcgc gtgggtgagg atgcgcgggt aggtgtcctg tttgacgagt 10440  
gcgggcaccg gtgccgggtc ctcgcccag gccgctcgt cggcggcgag gtcttcgcgg 10500  
ctggacagga acgcgcccaa gttcggcacc actggcacga actggttgtt cagcgcggtg 10560  
cgttcggggt ccggtcgcac gtcggcgtag aaccgcgccg aggtgcggga gttgggggcg 10620  
acgttggcga accagccgga gccgtcgggt tcgcgcagca ggttgtagct gagccagtcg 10680  
gtgtacttcc gcgccatcgt gacgaggttg gcgttggttg actggcgcca cgcgtcggcc 10740  
atctccggca gcatcacctc gaagttgtag ttggtgtccg cgccggtcaa ctcgtagaag 10800  
tagcccgaag gactctgccc gttcgcggca aggcgcgcga acgcctcggc caactgctgc 10860  
cgcagcgcgg ggtcggggtc ctggttcagc gcgagcgagg agccggcgag cccggcggtg 10920  
acctggttcg cgtagtggat tgtgggctgc caggcgctac cgttgcccgg gttgaggagc 10980  
cacgtcatcg ctttgcgag cgccgagggt atctgcgact ggcgttgcg caggatgttc 11040  
gccgcgcgca gcaggcggtt ggtcttgctc aggtacccca acccgaagcc ggtggcgcg 11100  
agcccggtgt cggtcggcga atactccggc cagagccgt catcgtgctg caacctagg 11160  
tagtgtccga gcccgcgctc gagcgcggcg agcagagtcg cgtcgccccg gtacgggttc 11220  
cacgtgcggg attgcgtggc gaaccaggcg agcgtgtaga cgtgctcctg gacgcgggcg 11280

ttgtaggaca cgcgcggt ggcgcaccag cctccggcga agaagcgcgt cgagtccatg 11340  
 tcggcgacca tcggcgagac cgcgggtgagg tacgacgcga atcgtctgctc ctcgggtgcg 11400  
 aacaggcgcc ggttcggtgt ggtggggcgc gggggcgggca gtgcgagtgc gcgggtcgcc 11460  
 agtgccggcga gcagtccgag agcggcgggc cgggtcatca ggctgcgacg gctgaacgta 11520  
 gtcacggggcc tacctccttg tggccgatca accctcaccc gctgcgtagc cgcacgtcaa 11580  
 gatgataatt cgaattatta tgggcttgac gacgcgtagg ccgacgacgc agaattcctg 11640  
 ccaattcgta ttggcaagcg ggggtgctcg tggcccgacg gctcacccag caggacatcg 11700  
 cccggatggc aggagtcagc caggccacgg tgtcgttgt gctcaacaac cggaaggacg 11760  
 gcaacgtccg gatcgcgcg gagacccgtg cgcaggtact ggaggtgatt cggaagaccg 11820  
 gctacgtcgc gaacccgatc gcccgacagga tgcgcgatcg gcacaaccgc atcctcggcg 11880  
 tgttcaccta cgaggcggtg ttcccagagca cccacgcgaa cttctaccag tcgttctctg 11940  
 aaggcatcga ggaacaggcc gaggaggtcg gctgtgacct gttgttgttc accagcgcca 12000  
 aggccaccgg ggagcgggcg cggattttcg gcgacgacag ccgggtgcg ctcgcccagc 12060  
 gcagtctgct gctcggtcgc acggtcgacc gcgacgacct gaccagctg ctcgcccga 12120  
 gcatcccgtg cgtctccatc ggacgacgcg acgacgcggg cgggtccggtg ccgacgctg 12180  
 gggccgacta ccgacccgcg gtgcgagacc tgggtggaccg cgcggtcgcg ctcggccacc 12240  
 gcgggttcgc gtacgtgggg tctggtgggg gcgcggagtc gtccgcggat cggctgcgag 12300  
 gcttcgcgga agccgttgcc gcacatggcg tccaagggat gcatgtggag accccacagc 12360  
 tcgatcagct gcgcgaagcg ggcgtcaccg ctgtgctcac cgaagaggtg tcggacgggg 12420  
 ccgcgctcgt gctcgcggg cgcgaacgcg ggctctccgt gccggggcgac ctcgcgctgc 12480  
 tctcgtcggg tgccgctacc cggtcggcac cggacgacga cgtggcgctc accggtttcc 12540  
 gcatccccag gcgcgagatg gggcgccggg cgggtgcaggc gctgaccgag gtgctcgaaa 12600  
 acggcaccac accgcaagaa ctgctcccgt gcgagttcgt cgagggtcg acgctggggc 12660  
 caccacgctt ttgaccagga ggaactgttg ctcgaccaca ccacggacgt tgcgtcggtt 12720  
 gggcgcggaac tcggcggtgt cgcgcgcga ctcgcgttg cgcgcgcggg ccgccgggtc 12780  
 gtgctcaccg aggagtacga ctggctcggc ggccagctga ccagccaggc cgtgcgcgcc 12840  
 gacgagcaca gctgggtgga gcgcttcggc gtcaccgcga gctaccgggc gttacgcgac 12900  
 ggcatccgcg actactaccg ccgccactac ccgctgaccc cgcgcgcacg ggctgtggcg 12960  
 gagctcaacc ccggtgcggg caacgtgagc aggtctgcc acgagccccg cgtcgcgctc 13020  
 gcggtgatcg acgagatgct ggcgccgttc cgcggcagtg gcaggctgac cgtgctgcag 13080  
 ccgtaccggc cgggtggcgc gcacaacgac ggcgaccgga tcgtgtcggt gaccgttgcg 13140  
 caccgcgaca ccggtgaaca gatcgagctc tccgcgccgt acatcctgga cgcgacggag 13200  
 acgggtgaac tgcttcggtt gtccagcacg gactacgtca ccggcttcga gtccactctg 13260  
 gacaccggcg agccgagtg gcccgacgtc gcgcagccg cgaacatgca ggcggtgtcg 13320  
 gtgtgcttcg tggtcgacca cgtcgacggc gaccacacca tcgacaaaacc ggcgcggtac 13380  
 gacttctggc gcgcgtacca gccggacttc tggggcgacc ggatgctgtc gttccgctcc 13440  
 cccaaccgcg gcacgctcgc gatctccgaa cgtacgttca cccgaaccc ggacgacgac 13500  
 ccgctcggcg tcgtgtcgga ccagcggtc agtgccgggtg acagcaatct gtggacgttc 13560  
 cggcgcatcg ccgcgcgtcg caacttcgtc gaggtgcct acgacagcga catctgcctg 13620  
 gtgaactggc cgatcatcga ctacttcgag tcgccggtga tcgacgtgcc ggacgccgac 13680  
 gcgcacatcg ccgcggcgcg ggaactctca cgttcgggtg tctactggct acagaccgag 13740  


---

 gcgcgcgcgc cagacggcg caccggcttc cccggcctcc gectgcgcgg cgacgtcacc 13800  
 ggacgcgcgg acggtctcgc gcaggcgccg tacatccgcg agtccaggcg catcagggcc 13860  
 gagcacacga tcgtcgaaca ggacctctcg ctcgcgcgtg gcggcgacaa ggggtcggtg 13920  
 cagcacgccg acgcggtggg tgcggcatg tacggtatcg acctgcaccc ctccaccggt 13980  
 ggcgacaaact acatcgacgt cgcgagctgc ccgttcgaga tcccgctcgg cgcgctgac 14040  
 ccgcaacggg tggagaacct gctaccgcg ggcaagaaca tcggcaccac ccacatcacc 14100  
 aacggttccc accggctgca ccagtcgag tggaacgtcg gcgaggtcgc gggcgcgctc 14160

gctgccttct	gcttggcgca	ccgagtcacc	cctcgcgcgg	tgcgcaatac	ccctggcctg	14220
ctcgcggact	tccagcagtg	tctggaacgc	gacggggctg	agctccgctg	gccggacgtg	14280
tccggctact	gacgcaggga	gacgaaaatg	acaaagctgt	cacgacgact	cacggcactc	14340
atcgtcgcag	ggctgttcgc	cctcaccggc	tgcggtggtg	gatcaaccgc	acagtccgga	14400
ccgaagtgcg	tgcgcatgac	cgtgtggact	gccaacgcgg	cgcattctca	gctgctcaac	14460
gacatcgccg	ccgagtacaa	ggcctcgcac	ccggacatcg	ccgagatcaa	gttcgactcg	14520
gtgcccgcgg	acggctacac	caccacgctc	accacccaga	tgcgcggcgg	taacgcgcgg	14580
gacctggcct	ggatcctgga	ggagtccgca	ccggacttcg	tggcgtccgg	tgcgctcgcc	14640
ccggtgcgcg	gcaagatcga	gaaggccgac	gagctcgtgc	cgtccgcgac	gaagctgtgg	14700
gagaaggacg	gcgaactgta	cgcctaccgg	ttctccacct	cgccgttcgg	cttgttcgtc	14760
aacaccgacc	tgggtgaagg	cgctcggcg	gactggacct	gggaccaggc	gatcgcggt	14820
gcctctgcgt	cggcgccgcg	ctccggcaag	ggcggcctgg	tactgccgga	cttcaagtac	14880
cagaactggg	cagtgtgttc	ctctatctgg	cgcggctggg	gagctgatgc	gtggagcgcg	14940
gacggtcgct	cgtgcgggtt	ctccagcagc	gagatgaacg	acgcgatgtc	cttcttgcac	15000
aaggccatct	tcaccgacaa	ggcgattccg	ggccccggca	cgacggtgga	cttcttcgcc	15060
ggcgacgcgg	cgatggcgat	cggccagatc	tcccggtcca	gtgcgttgaa	ggacgcgaag	15120
ttcggctgga	cgctgtgtcc	gctgccggcc	ggtcgaagg	gtgactacgc	ggtgatcggg	15180
caggccggga	tcggtgtgct	gaagcagtcg	cacaacgtcg	acgccgcgac	ggacttcctc	15240
gccttctca	ccaaccagac	caactccgcc	aagctcgcgc	agttcttccc	ccggcgcgcg	15300
tcgtcgtgc	tcaacgcgga	gacgctcgcc	aagagcaacc	cgttgatcaa	ggccgagcaa	15360
ctgcagtccg	tcgtcgtcga	cggcatcaac	aagggcgtcg	tgaagccgag	ccacaagggg	15420
caggaggagc	tgaaccagac	gatccgcgcc	gcgctcgacc	cgttgtggaa	gccggacgcg	15480
aacgtgcaga	acgtgttgaa	cgacgtgtgc	accaagatca	aaccgctgct	ggagaacaag	15540
tgacggcggg	cgcacactcc	acgcgggccg	tagcccgcaa	ggggacgtcg	tactggacgc	15600
agcgcaggcg	cgacaatctg	gttggttacc	tgctcgtcgc	gcccgcgctg	ctcggcagca	15660
tcgcgttcgt	gctggtgccg	ctggccttgg	tcggctggta	cagcctcaac	gagtggaaac	15720
tgctcgcggg	cacgttcgag	ttcgtcggcg	cgcagaacta	ccaacagctg	ctcgccgacg	15780
agaagctgcg	cgactcgtcg	gtggcgacca	cttggttcgc	cgcgggcctg	gtcgtgctca	15840
acctgtcact	agcgtgtgtg	ctggccgtgc	tgctcaacca	gaagctgagc	ggcaccacgg	15900
tcttccgcac	gttgtttctc	tctcccgtcc	tcgtgtcgct	ggtggcggtg	acgctggtgt	15960
ggcagctgat	actgcagccg	gagggcagcg	tcaacggact	gctcgggttc	ttcggtgccg	16020
acgggccgaa	ctggctgcgc	ggtgagtcga	ccgcgatggg	ctcggtgatc	gtcgtgcagg	16080
tgctcaagaa	cgtcggcctg	aacatggtgc	tgttcctcgc	ggcgctgcaa	ggcgtgccgc	16140
agccgttgta	cgaggcggcg	aagatcgacg	gggcgggtgc	ctggaccggg	ttccgccgca	16200
tcaccttgcc	gttgatcagc	ccgacgatcc	tgctcacgtc	gatcatcacc	atcgtcgggt	16260
cgtgcaggt	gttcgcgcag	atcgcgggtg	tcacgcaggg	cggtcggggc	acgtccacga	16320
ccgtactgat	ctactacctg	taccagcagg	cgttccagtt	ccaccacttc	ggctacggcg	16380
cgacgatctc	ggtactgctg	ttcgtcatcg	tcgccgcact	caccctgctg	cagtggcaga	16440
tgcgccggaa	gtgggtgctg	catgaggctt	agggtcaaga	tcacgctata	cgggctgctg	16500
tgctcgtgt	gcgtgccttt	cgtgttccca	acctggtgga	tgatcacctc	gtcgtggaag	16560
cccatcagcg	agatcttctc	cacctcgccg	ctgccgtcgg	agtggacgtt	ctccacctac	16620
<del>cggeaggtgt</del>	<del>tcgagatgca</del>	<del>gcggttcggg</del>	<del>cagcagtaact</del>	<del>ggaacagcct</del>	<del>ctacatcgcg</del>	<del>16680</del>
gtgatcgtca	cgacaggcac	aatggcggtg	gcggcgatgg	ccggatacgc	gttcgcgcgt	16740
atccggttcc	cgggccagaa	cgtgctgttc	gtggtcgtcc	tgatcggcct	gctcatcccg	16800
agcgaggta	ccatcgtgcc	gctgttcaag	atgttccagt	cgtcggcct	gaccaacacg	16860
cactggccgc	taatcatcgt	gccgatcttc	ggggcgccct	gtgtgctggc	gatcttcac	16920
atgcgccagt	tcttcacgcg	cctgccgtcg	gagctggagg	aggccgcgcg	gatggacggc	16980
ctcggacgag	cgggcatctt	ctaccgggtg	gccctgccgc	tgtcccgggc	cgcgctgggc	17040

gcccgtcgcga tcttcacg cctgcactcg tggaaacctct acctgggccc gatcgtctat 17100  
ctgtccacac cggacatgta caggttgccg caagcgctca cgcagttcgt cgacaactac 17160  
ggcggccccga tgtggaacgt gcagctcgcc gccgcgacca tgaccgcagt gccgattctc 17220  
gtcgtgttcg tgctggcgca gcggcagttc atcgagggac tcgcgcacac cgggggtaag 17280  
gggtgagtc cgcgatcgac ccggagatcg ccgcgctggt agacctcgcc agaccaccgg 17340  
tgacgccagc gggaatcgac gcggtccggg cgggaggacg cgtcgtcacc gatgccgagc 17400  
tgacgcgcgg tggcaccgtg acgttcgccc atgccgatgc cgatggcgtg cccgtgcttg 17460  
tgctgcgccc agcgggtgtg ccacgtcttc cgggtgctga cctgcacggc ggccggatgg 17520  
tcgcgggcac ccggcgacc gacctgcacg tgctggccga gtgggtgtcg gagctgggtg 17580  
tcgtgctggt gtcgccggag taccggtcg cacccgagca cccgcacccg gcgcgctcg 17640  
aggactgctt ccgggtgctg gagtggatgt cccgcaacgg cttcggcccc cctgtcgttg 17700  
cgggtacatc ggccggcgcc gggctggcg ctgcggtgac gctgatggcc cgcgacctcg 17760  
gcggtccgcc gatcttgccg tcggcgcgca actgctcgaa cgcgtccagc gtctccgcca 17820  
acggcgtgct accgcgccag tgcagcaggt agacgtccag gtgatcgggt cccaaccgcc 17880  
gcagactgcg ttcgcacgcc tcgacggcgc caccgccacc cgcgttgctg ggatacacct 17940  
tgctgaccag gaaaacctcg tcgcggcgcc cggcgatcgc ggcaccgacg acctcctcgg 18000  
cgcccccgct gccgtacatc tcggcggtgt cgatcagccc caggccgagg tccaggccgc 18060  
gccgcagcgc ggcgacctcc gccgcgcgt gccccacgcg ctcccccatg ccccggtgc 18120  
cctgccccag caccggcaat tctgccccgc caccgagttg aaggcctcgc atgcacactc 18180  
cgtttcagtc agcgtgctg cggcagcacg gccaccagct cgaagccacc gtctcggac 18240  
ggcccgggcg ccaaccgccc accggccatc tgcaccgct ccgacaggcc ggccagccc 18300  
gaccgcgcgc tgcgacctc gatggcgggt cccgcgcggt cgggggcccgt gtttgtgacg 18360  
gcgacgtgca gcgagccctc gcggaacacc agccgcacct tgatgtcagc accgggtgcg 18420  
tgcttgaag cgttcgtcaa cgctcctgg acgaccgggt aagccgcgcg ctcaaccgca 18480  
ggcgacaccc cgtgcgggtc gcggtatctc gaggaaagct ccacatccac accggcgggc 18540  
cgcgcccga cgaccagctc cgggatcgtg ctcaaccggg cccgctgcct gggggcgccc 18600  
agcccgagcg tcgcgcgcac ctgcctcagc gcctgcttcg ccaaccgcgc cagccgccc 18660  
gcggtctcct tcgcaccggg atcctcgggt gtggcgcca acgcgcgcga ctgcaccgcg 18720  
attaaggtca cgtgatgtcc caccgcgtcg tggatctccc gcgcgatccg ggcccgctcc 18780  
tcgcgcgcgc cgtctcggc cttgcctgc aactcgcct cgggtggcgc ctccagcttc 18840  
cgcgactgt ccgccagctg ctcccgcaac gccaccagcg ccccaatgc cgtcggcgcg 18900  
cctgcccga gcagcacgaa cgccacggtc aggatgatcg acccgacccc cagcgactgc 18960  
ctgatcatca ccggtgccac ggcgaccgcc gcggtcagca acaccagcc gaccaacatc 19020  
gggatccgcg gctgcgcgcg ccccaacgcg tacaacgcca cgaccgtcgg cggccaaccc 19080  
aaaccaccgg ccaacgctgg catgcacagc agcaccgcga tccgcggcca ccgccgaac 19140  
ggcagcagcg ccgaagccac caaacggac accgtcgcac agccgaacgg gtcgggcgcg 19200  
gccagcacag ctgccgcgg aatcccgcag accagcgct caacgcgcg ccgagaccag 19260  
cgcaccgaac tcaccggaag ccgcccgcca cgcgctgcgc gatcaacgcg gcctgcaccc 19320  
ggttctgcgc gccgagcttg ctcaacaccg tcgacacgta gctcttgacc gtgcctcgg 19380  
tcagcccgag ccggacccc atatcggcgt tggaccgacc atccgccaac aactggagca 19440  
cctgccgctc ccgcccgga aacccctgca ccagctgcgt ctgcgtgcgc ttgtcggaca 19500  
gactgcggaa cctgggcaac aaccgcgcag tgatccgcg atccagcacc gcccaccgg 19560  
ccgccagatc gtgcaccgcc cgaaccagca ccgcgggctc cgcgtccttg agcaggaacc 19620  
cgtgcgcccc gaaccgcaag gcacggcca cgtaatcgtc caggtcgaac gtcgtcagca 19680  
ccgccggcac cggcggtatc gccaacgtcg ccaagtcccg caacgcgcgt agcccgctc 19740  
tgccgggcat ctgcacgtcg atgaggcgca cgtcgacgcg gtacgcgcgc acccgggcca 19800  
gcaactcatc cccgttgccc gcctcggcga ccaccggat cgagccgtcg cctccaaca 19860  
acaccgcag gccacggcg agcatcgct cgtcatcggc aagcaaacg cggatcggcc 19920



ccccgcgctc	gtcagcccg	acctctcggt	cacctccctg	ttccgaggg	cgccagcaga	19980
ctagcggggc	cgcgaaagcg	cagcgacac	gccacccgag	cgcactgtgt	cagcggcgcc	20040
cgcccttggc	cttccgctcc	cgcacccgca	ccgagatccg	caccgggcta	cccacgaacc	20100
cgaacgtctc	ccggaacctg	cgctccagga	accgccggta	cccggcctcc	aagaacctcg	20160
tgggtgaacag	cacgatcgtc	ggcgggcccg	actgcgcctg	ggtggcggaac	aggatcttcg	20220
gctgcttgcc	gccccgcacc	ggcgggcggt	tggccgcgac	cagctcgga	agccaacctg	20280
tcaaccggcc	ggtggagatc	cggtgtccc	acgaatccag	cgccgtg	agcgtcg	20340
ccagtttgc	caccgcgga	ccgtcttcg	ccgacacgtt	gaccgctcg	gccaccgca	20400
cccgaccag	ctcgcggtcg	atctccttct	ccaactgggtg	gcggcggtcc	tcgtcgacca	20460
ggtcccactt	gtttagggcg	atcaccaatg	cccgaccggc	ctcgacgacc	atagtgatca	20520
cccgacggtc	ctgctcgctc	aacggctcgg	agccgtcgat	gagcacgac	accacctcgg	20580
ccgctcgat	cgcggccttg	gtgcgcagcg	acgcgtagta	ctcggtg	cttgcggtct	20640
tcacgcgctt	gcgcagccc	gcggtgtcga	cgaaccgcc	cacctggc	tccagctcca	20700
ccagcgagtc	caccgggtcc	accgtggtgc	ccgcgacgtc	gtgcaccacc	gaccgctctt	20760
ccccggtcag	cttggttcagc	aggctcgact	tgcccacgtt	cggttgccg	accaatgcca	20820
ccgcgcg	cccaccggtc	gcgcgaaga	tctcgcg	cgtctcggg	aacacctcca	20880
ggacggcgtc	cagcaggtca	ccggaaccgc	gcccggtgaa	cgcgctgacc	ggcatcggt	20940
cgcccaacc	cagcgaccac	agcgaatgca	cgtcggagac	gcctcgctgg	tcgtcgacct	21000
tgttgggc	cagcagcacc	ggacgcttg	accgccgag	cacctggcc	acggcttct	21060
cggctctcgg	ggcaccgacc	cgggcgctca	ccacgagcag	caccgcatcg	gcggtgtgca	21120
tggccagttc	cgctcgcg	gccaccgacg	cctgcagtcc	cttcgcgtcc	ggctcccagc	21180
caccggtgtc	gaccaccgtg	aagcggcgcc	cgttcacaa	cgctcgtag	gccaccgat	21240
cccggtgac	accaggggtg	tcctgcacga	cggttccc	acggcccagc	agccggttca	21300
ccagcgtcga	cttgcccacg	ttgggacggc	cgaccacggc	cagcaccggc	tgcggttcg	21360
tcggtctctc	gccctcgacc	gcttcgacct	cgtcgaactc	agccactca	gcctcgctcg	21420
accacgtgcc	gtccaggcct	tccaccgact	cgtcggtcac	gctcgtttcc	ctctcgaata	21480
gtcttggtc	tggtcgcga	ccagccgatc	cagctcgcg	atcagctcga	ccagttcggt	21540
gcgcaccgcc	tcggtcgcca	cgaccagccc	ggcgcgcc	ttgcccgcg	gcagcacgat	21600
cggctcccc	aacaacacgt	cgatcctcg	tagcagccta	cggcccggg	catccggcg	21660
gcgcgtgccc	cggcaggcca	ccggcaggat	cagcgccctc	gacgtccg	ccaaccacgc	21720
cgcaccggtc	tgcgcattgg	tcacatcgcc	atcaccccg	gtgccctcc	ggaacacccc	21780
gacgagacca	ccggcccgc	gcaccgcac	cgccgcagc	agcggggcg	ggtccggcg	21840
accgcggcg	accccgatct	gcccgatgc	gcgcaggaa	cagcccagcg	caccgcggaa	21900
catctcctgc	ttgatcagga	acaccgcgtg	gcgcgggatc	attccgaaca	gcagcgccc	21960
gtccatcatc	gaactgtggt	tggccaccag	taccaccgg	cccgtgcg	ggatccgctc	22020
cgcgccgtgg	atccgcaccc	ggtacggcag	ctgcacgatc	cgcgcgaaa	tccactggcc	22080
gcaccgggtc	attccaccgg	acgcgcctc	gggcagggtc	tcgtcgctca	tggcgctgc	22140
cgctccgcg	cgaccagcag	gccacggcac	tccacgtgct	tgcgcaactg	cgtgagcacc	22200
tccaccacac	ccaggtccgt	ggtgtcgagc	tccaccgcgt	cgtcggcctt	gcgcagcggc	22260
gccaccttc	gccccgagtc	cagcgcatcg	cggcgtgca	cgtcgcgtg	ggtccggtcc	22320
aggtcaccgg	ccgctccctc	ggagacgtcc	tgcgcgtgc	ggcgctg	gcggcggtgc	22380
geggacgcg	tcaggtagac	cttcaaacc	gcatecg	ccaccacgtt	gcccgtg	22440
cgctccctcga	cgaccatgcc	gcctggcgac	accagcgct	caccgatcag	cttcgctgg	22500
tgggcgacca	actgctccc	cacctcg	accgccgaca	ccgccgacac	cgcgccggtc	22560
acctccggcc	cgcggtatctc	ccgcccacg	tccgagccgt	ccaggaagat	ctcagggcc	22620
tccggatcg	tgccaccgt	cagccgcgt	gcgcgacca	cgctgccac	tgcggccgga	22680
tccgacgggt	cgacctcg	gcgcagcacc	gccagcgtca	cggcgcggt	catcgctcc	22740
gtgtcgaggt	aggtagcccc	caatgcgaac	gccaacttgc	gggacaccgt	cgacttgcc	22800

gtgcccgaag ggccgtcgcgccaccaca ccaaggagcc tggcgtgtgc cacgtgcatg 22860  
 tccttcccga gccggaaccg atcgccccga caccggccg gaccaaccgt ccattctgcc 22920  
 tggtcgctc ctgccgatcc caatcgcatc cagatccaat ccggaactgc ggacaaatac 22980  
 gaaagtctta cgaataaggt cttgggtgtg gatgtaactg tgacccctt gccagggtc 23040  
 ggcacgcagc aggacttcac caccgatcc ggccaccgca tcggcgtgat cacctaccg 23100  
 gacggccgct tcgagctgat cgtctccgac cacgaagacc cagacaaggt ggccgcctcg 23160  
 gtgccttga ccacctccga gaccagcacg ctggccaacc tgctgggcgc ccgcagctg 23220  
 gtgcccagc tcaacgagca gcacagcgag gtgcgcggga tcaccacctg gcagctgtcc 23280  
 gtgcccccg gctcccccta cgccggccga accctcggcg acacggagat gcgcaccgc 23340  
 acctcgtgt cgatcgtcg ggtcgtccg gacggcaccg tgacccctc accgcggccg 23400  
 gacttccagt tctccgccg cgacctggtg gtggtcgtg gcaccgcgaa ggggctacg 23460  
 gcagccagcg aaatcctgga aaagggctga tccggcccg cccgaacact tccccggcca 23520  
 cctgaccggg acacgcgtca ccaacggcg tccagcgct gccgatcccc taacaactcc 23580  
 atacgcaggg cacacgcgag aggaacacg ttgcaagaca cggcgatctc cctaatagaa 23640  
 ctgggtgcg ttttcttcg tctgggcac ctcgggagac tcgcatggaa aatcggggtg 23700  
 tcaccgatcc cgctctacct gatcgccgg ctggccttcg gcaccggtg cctggtcccc 23760  
 ctgcacggca tcgaaccgtt caccacctc gcctccgaga tcggcgtggt cctgctgttg 23820  
 ctgctgttag gcctggaata ctccgcggt gaactcgtca ccggcttac cgctcctgg 23880  
 ctggccggtg tcatagacat cgtgctgaac gcggtaccg gtgcgctggt ggccctgctg 23940  
 ctgggttggg gtccgctcgg cgcttcaca atggccggcg tcacctacat ctcttctcc 24000  
 ggcacatcg cgaaggtgct cggagacctg ggtcggctcg gcaaccggga aacgcgggtg 24060  
 atcctgtcga tctgtgtctt cgaagacctg gcgatggcg tctacctgcc gatcctgacc 24120  
 gccgtgctgg ccggggtcag ctctctcggc ggactgaccg cggtcggggt gtgctgggtc 24180  
 gtgatcacc tgggtgctgg ggtcgcgct cgcttcggca agtacgtgtc cgcgctgggtc 24240  
 gacagcccc acccgagggt ctctctgct cgctgctgg gggcgcgct gctcgtcgcg 24300  
 gggatcgct cggagctgca ggtctcggcg gcggtcggct cgctcctgct gggcatcgcg 24360  
 atctccgggt cgaccgcggc gaacgcgacc cgcgtgctgg agccgctgcg cgacctgttc 24420  
 gccgcgctgt tcttcgtcgt gtccggctc aacaccgatc cgagccagat cccgccggtg 24480  
 ctgcgggtg cgctgctgct ggcggtggcg acctgcgca cgaaagtgt caccggtg 24540  
 ttcgcccgc ggatgcagg tgcggccg atgggtcggc tgcgggcccg tgccgcgctg 24600  
 gtggcgcg cggaattct catcgtgat gccgggctgg ccgtgcctc cggcgcggtg 24660  
 ccgggcgagc tcgcccct cgccaccgc tacgtgctg tgatggcgat cgtcggcccg 24720  
 gtgcggccc gcgtcgtcga gccggtggcc cggatgttcc tccccagcca ggcgaagaag 24780  
 cctgaccgg aaaagcctcc ggctgtccc agcggacggg ccggaggctt tctcaacgct 24840  
 gcgataacg cgggcgggtc tggcaaacgc aacaccggtg cgcgaacagc ctgagaggcc 24900  
 cactgcgcg tagagcgagc cgatctcct gcggttcaac ggccggatcg cgcggggcg 24960  
 ctggttggt agccgcacat cgccgatcc cgtccgcacc aggcgctgca ccgggtagcc 25020  
 gacgtgctt agcagccgc gcacgatgc gttgcggccc tcgtgcaaca cgatctccac 25080  
 cagcgaccg ttctggttca tgcgaccag cttgaaccg tccaccttca ccagaccgtc 25140  
 ctccagctcc acgcctcgc gcagctgct cccgaggtcc ttcgggatcg acccagcac 25200  
 ctccgccagg tacgtcttgc gcaccttga cgacgggtgc atcagccgt tcgccagctc 25260  
~~geegtegtt gtgateagea geaggeete ggtgteetgg teaggeee eaeegtggaa 25320~~  
 cagcttgccc tcccgttccc gcaggtaatc accgatgcac ggacgcccct tgcgtccga 25380  
 catggtgcac aggatgcgc gcggttgtt cagcagcagg tgcgtcagat cgtcgttgac 25440  
 catcaccgg ttgcgtcca cgtggaccac cgcggtcgtc gggtcgacc ggcccccag 25500  
 ctggtgacg acctcgccgt cgacctgat cgcacctcg acgatcatc ctcggccgc 25560  
 gcgcgcgag gccaccgcg cctgggaaa cactttctgc agccgcacgc cttccgaatc 25620  
 ggcagtgtc ctgtacggg aacggtgctc agacgtcatc aatcgaatcc acttcgggca 25680

acaagggggc	gagcggcggc	aactccttca	acgacgacag	ccccagccgc	tccaggaaca	25740
gctcggtcgt	gcagtacagg	atgccgcccc	tctccgggtc	ggttccggcc	tcttcgataa	25800
ggccgcgccc	caccagggtta	cggatgacgc	cgtcgacgtt	cacaccccgt	actgccgcga	25860
cccgcgaacg	cgtcaccggc	tgccgggtacg	cgatcaccgc	gagggctctc	agcgcggcac	25920
gcgtcagctt	cgaacgctgc	ccgtccaaca	ggtaccgctc	cacgtacggg	gcgtacttct	25980
cccgcgtgta	gaaccgccag	ccgtcgccga	cccgcgcgag	gtcaacgccc	cgttccgcct	26040
cggcgtagcc	gtccgagagc	cgccgcagcg	cggaccggat	ccgggccacc	ggctgctcca	26100
gggtgtcggc	cagcagctcc	tcgccggccg	gcacgtcgac	cacgagcagc	aacgcctcca	26160
gcgccgcctc	cagcgccgaa	tcggagggtca	ggtccggcgc	actcggcgcc	tccaccgcct	26220
cggaatcccc	agcggccgcc	gacgggtcca	acgggttctg	cggctcagcg	ggcggcgggg	26280
tcggcgactc	ctccgcgacc	tcgcccgcg	cctccggctc	cagctccgtc	acccgtactc	26340
ctcttcttca	tggctggcgc	gctgcgcctc	ggcttcggcc	tccgcctgcg	ccacgctgcc	26400
gcccacccag	ctcacgatca	gctcaccag	cgggggtctcc	tgctcgaaca	gcaggacctt	26460
ctcccggtag	agctcgagca	gcgccaggaa	ccgcgccacc	acctcgacgg	tgtgcccgca	26520
gtccgacacc	agctccgaga	acgtcgcgct	gcccttctcc	gccaaaccga	cccgcagcag	26580
cgcgcgctgc	tcgcgcaccg	aaacgccgtg	ctgggtggatg	tgggtccagcg	aaacggctcg	26640
cgggtggcttc	ggccggaaca	ccgcggccgc	gatctcggcg	aaccgctgcg	gcggaacccc	26700
gatcaccact	tccggcagca	ggttcgcgaa	ccgctcctcg	accgacaccg	aacgcggata	26760
gcgccgcagc	gcccgggcct	ccaactcgcc	gaacagcgcg	gccacctgct	tgtacgcgcg	26820
gtactgcagc	agccgcgcga	agagcaggtc	gcgggcctcc	agcagcgcca	ggtccgcctc	26880
gtcctcgacc	tccgcggccg	gcagcagccg	ggccgccttc	aggtctaaca	gcgtcgcggc	26940
gaccaccagg	aactcgggtg	tttcggttcag	gtcccaccgc	tcgccgagcg	ccttggtgta	27000
ggcgtatgaac	tcgtcgggtga	ccttggtgcag	cgctacttcg	gtaacgtcca	gctggtgctg	27060
cagatcagc	tgacgcagca	ggtcgaagg	gccctcgaag	ttgtccagcc	gcaccgtgaa	27120
ccggccgcct	gcggcgggat	cctcgctcga	ctgcgcgcgc	ggtaccgcgg	ggtctgtcac	27180
tctccgggct	cctgcaggcg	ccgcaccagc	accgagtcgc	cgccgttctt	gtcgaggtcg	27240
gcgagcacca	cggcaaccgc	ttcgcgcacc	acgcggcctc	ggtccacggc	gatgccgtgg	27300
tcggccccga	atgccagccg	cgcctgctcc	aagcccagca	actcctcgtc	ggacacgtac	27360
acgggtgatct	tcgaatcgtg	cttgccgcct	ccagagccgc	cccggcccgc	cgcggtgccc	27420
gcctgctccg	ggtccggcgg	attcgcgtga	tccgtccgag	gaacagtagt	acggaacaat	27480
tcggccgctc	cgggcagggc	gaccgcagcg	gtcatcgggc	gatcacctcg	cgcgccagt	27540
cgcggtacgc	ctgcgccccg	gccgagcggg	gcgccagcg	ggtgatcggc	tcgccggcca	27600
tcgtggtctc	cggaaaccgc	accgtgcggt	tgatcaccgc	atcgaagacg	atgtcgccga	27660
tcgcctcgac	gacgcgcgc	atcacctcgc	gcgagtgcag	cgtgcgcggg	tcgaacatgg	27720
tcgccaggat	gccgctgac	tccagcctcg	ggttcagccg	ctccgcacc	ttctcgatgg	27780
tgctgatcag	cagcgccacc	ccgcgcaggc	tgaagaactc	gcactccagc	gggattatca	27840
cggcgtccgc	ggcggccaac	gcgttcaccg	tcagcagccc	cagcgacggc	tggcagtcca	27900
ccagcacgta	gtcgtactcc	gccagcgccg	gctgcagcac	ccggtgcagc	gtctgctccc	27960
ggccccacct	cgcgaccagc	tgacacctcg	ccgcggacag	gtcgatggtg	ctgggcagca	28020
gggtccatgcc	ctcgacgggtg	gtctgcccga	ccacgtcatg	gatggcgacc	ggccgttcca	28080
tgatcacgtt	gtagatggtc	tggtcgagct	ggtgcggctg	gaccccgagg	ccgaccgaga	28140
<del>gcgcgccttc</del>	<del>cggtatcgaa</del>	<del>tccaccagca</del>	<del>gcacccggcg</del>	<del>ccggtatctc</del>	<del>gcgcgccttc</del>	<del>28200</del>
cgcgcgaggtt	gatcgtcgac	gtggtcttcc	cgacgcgcgc	cttctggttg	cacatcgcca	28260
gcaccgatgc	gggaccgtgc	cgggtccagca	gcggcggctc	ggggatgtgc	cgacgtggcc	28320
ggccagtggg	accgaggccg	cgggggttcg	gcgccaggcc	gacctcctcc	tcgcggcccg	28380
ccgtctcccc	ggcgatgctc	aggtcgaccg	cgccccggga	ccgccttctg	gcggcgggct	28440
gcggcagcga	catggcgatc	tgactccttt	tggtgccggt	ggcgatcaaa	cgcagcctaa	28500
gtggcattcg	ttaccagcgg	caacgcgcct	cgccgtgggtg	tcggaagtcc	cctgtgatct	28560

gcgtgtgata aacgtctgc gcgtacacca cccctcgcga ttcgcaggt caccaagtca 28620  
 ctccgtagcg tgtgcgcgcg gatgcgtcgt cgcgtagacc tcccgcagcg tgttcaccgt 28680  
 gatcaaggtg tagacctggg tcgtcgtgac cgaggcgtgg ccagcagtt cctgcaccac 28740  
 gcggacatcc gcgccgccct ccacaaggtg cgtggcgaac gaatgccgca gcacgtgcgg 28800  
 cgaaaccgag ccgttgatcc cggcgcgcgt cgcggcgggtc ttcaacgcgt tccacgagct 28860  
 ctgccgggaa agccgggtac cgcgcgaatt caggaacacc gcggcaactac cccgcccgcg 28920  
 cgcggccagg ccgggcctgg cccgtaccag gtacgcgtcc agcgcgcgca gcgcgcgtcg 28980  
 gccgatgggc accagccgct ggccggccgc cttcccttcc agcagcacgg tccggttggg 29040  
 gccgtcgatg tcgtcgaggt ccaaccgcac cgcttcggag atccgcgcgc ccgtcgaata 29100  
 cagcacttcc agcagcgcgc ggtcccgcag gccccgggca tccctgcgcc cggcgtgata 29160  
 gagcagcttc ccgacgtgcg cgatggaaaag cgccttcggc agccgccttc gcgggctcgg 29220  
 cggcgcgaca tcgcgcgcta cgtcgaccgg cagcagcccc tcggcgtgcg cgaagcgggtg 29280  
 caggccgcgg gcggccacca aagctcgggc cgacgacgac ggccgcagcg cgggcccgtg 29340  
 cccggtgcct tcccggagct ccgcgaagaa cccggacagg tgctccgaac gcacctcgcc 29400  
 gagcccagaa atccctgcgc cgaccaggaa ttcggcatac cggcgcaggt ccctggcgta 29460  
 ggagtcgaga gtgctgcggg cggtaaccgc ctcgacggca agtggtcga ggtaccgggt 29520  
 gatcgccccg cgcaggtccg gcggcagctc ctcgaaaccc ctcaccgggc gggaagcgc 29580  
 cggggccggg ccggccattc ggcattccgc gggcgaggtt ccaccgggc ctcgcggacc 29640  
 gcgtgggcgg cgagcagccc cgccaccccg gggccggtga cgatctcccc ggccagcgcc 29700  
 atccgcaccg cctcggcgag cgggaagcgg cggatc 29736

<210> 7

<211> 828

<212> DNA

<213> Saccharopolyspora spinosa

<220>

<221> CDS

<222> (1)..(825)

<223> ORF1; O-Methyltransferase

<400> 7

atg ttg cca ggt ggc gca cca aca tcg cag cag gtt ggg cag atg tat 48  
 Val Leu Pro Gly Gly Ala Pro Thr Ser Gln Gln Val Gly Gln Met Tyr  
 1 5 10 15

gac ctg gtc acg ccg ttg ctg aac tcg gtc gcg ggc ggc ccc tgc gcc 96  
 Asp Leu Val Thr Pro Leu Leu Asn Ser Val Ala Gly Gly Pro Cys Ala  
 20 25 30

~~atc cac cac ggc tac tgg gag aac gac ggg cgg gct tcc tgg cag cag 144~~  
 Ile His His Gly Tyr Trp Glu Asn Asp Gly Arg Ala Ser Trp Gln Gln  
 35 40 45

gcc gcc gac cgg ctc acc gac ctt gtc gcc gaa cgg acc gtg ctc gat 192  
 Ala Ala Asp Arg Leu Thr Asp Leu Val Ala Glu Arg Thr Val Leu Asp  
 50 55 60

ggc ggc gtt cga ctg ctc gat gtg ggg tgc ggt acc gga caa cca gcg 240  
 Gly Gly Val Arg Leu Leu Asp Val Gly Cys Gly Thr Gly Gln Pro Ala  
 65 70 75 80

ctg cgc gtc gcg cgc gac aac gcg atc cag atc acc ggc atc acc gtc 288  
 Leu Arg Val Ala Arg Asp Asn Ala Ile Gln Ile Thr Gly Ile Thr Val  
 85 90 95

agc cag gtg caa gtg gcc atc gcc gct gat tgc gca cgc gaa cgc gga 336  
 Ser Gln Val Gln Val Ala Ile Ala Ala Asp Cys Ala Arg Glu Arg Gly  
 100 105 110

cta agc cac cgg gtg gac ttc tcg tgc gtc gat gcc atg tcc ctg ccg 384  
 Leu Ser His Arg Val Asp Phe Ser Cys Val Asp Ala Met Ser Leu Pro  
 115 120 125

tac ccg gac aat gct ttc gac gcc gcc tgg gcc atg cag tcg ctg ttg 432  
 Tyr Pro Asp Asn Ala Phe Asp Ala Ala Trp Ala Met Gln Ser Leu Leu  
 130 135 140

gag atg tcc gaa ccg gac cgt gcc atc cgg gaa atc ctt cga gta ctc 480  
 Glu Met Ser Glu Pro Asp Arg Ala Ile Arg Glu Ile Leu Arg Val Leu  
 145 150 155 160

aaa ccc ggt ggc atc ctc ggc gtc acc gag gtc gtc aaa cga gaa gcg 528  
 Lys Pro Gly Gly Ile Leu Gly Val Thr Glu Val Val Lys Arg Glu Ala  
 165 170 175

ggc ggc ggg atg ccg gtg tcc ggg gac agg tgg ccg acc ggc ctt ccg 576  
 Gly Gly Gly Met Pro Val Ser Gly Asp Arg Trp Pro Thr Gly Leu Arg  
 180 185 190

atc tgc ctg gct gag caa ctt ctg gaa tcg ctg cgt gca gcg ggg ttc 624  
 Ile Cys Leu Ala Glu Gln Leu Leu Glu Ser Leu Arg Ala Ala Gly Phe  
 195 200 205

gag atc ctc gat tgg gag gac gtg tcg tcg agg acc ccg tac ttc atg 672  
 Glu Ile Leu Asp Trp Glu Asp Val Ser Ser Arg Thr Arg Tyr Phe Met  
 210 215 220

~~ccg cag ttc gcc gaa gag ctc gct gcg cac cag cac ggg atc gcg gac 720~~  
~~Pro Gln Phe Ala Glu Glu Leu Ala Ala His Gln His Gly Ile Ala Asp~~  
~~225 230 235 240~~

agg tac ggg ccg gct gtc gcc ggc tgg gcc gcc gcg gtc tgc gat tat 768  
 Arg Tyr Gly Pro Ala Val Ala Gly Trp Ala Ala Ala Val Cys Asp Tyr  
 245 250 255

gag aaa tat gcc cac gac atg ggc tat gcg att ctg acg gcg cgg aag 816  
 Glu Lys Tyr Ala His Asp Met Gly Tyr Ala Ile Leu Thr Ala Arg Lys  
                   260                  265                  270

ccg gtc ggc tga 828  
 Pro Val Gly  
                   275

<210> 8  
 <211> 275  
 <212> PRT  
 <213> Saccharopolyspora spinosa

<400> 8  
 Val Leu Pro Gly Gly Ala Pro Thr Ser Gln Gln Val Gly Gln Met Tyr  
           1                  5                  10                  15

Asp Leu Val Thr Pro Leu Leu Asn Ser Val Ala Gly Gly Pro Cys Ala  
                   20                  25                  30

Ile His His Gly Tyr Trp Glu Asn Asp Gly Arg Ala Ser Trp Gln Gln  
                   35                  40                  45

Ala Ala Asp Arg Leu Thr Asp Leu Val Ala Glu Arg Thr Val Leu Asp  
                   50                  55                  60

Gly Gly Val Arg Leu Leu Asp Val Gly Cys Gly Thr Gly Gln Pro Ala  
           65                  70                  75                  80

Leu Arg Val Ala Arg Asp Asn Ala Ile Gln Ile Thr Gly Ile Thr Val  
                   85                  90                  95

Ser Gln Val Gln Val Ala Ile Ala Ala Asp Cys Ala Arg Glu Arg Gly  
                   100                  105                  110

Leu Ser His Arg Val Asp Phe Ser Cys Val Asp Ala Met Ser Leu Pro  
                   115                  120                  125

Tyr Pro Asp Asn Ala Phe Asp Ala Ala Trp Ala Met Gln Ser Leu Leu  
           130                  135                  140

---

Glu Met Ser Glu Pro Asp Arg Ala Ile Arg Glu Ile Leu Arg Val Leu  
           145                  150                  155                  160

Lys Pro Gly Gly Ile Leu Gly Val Thr Glu Val Val Lys Arg Glu Ala  
                   165                  170                  175

Gly Gly Gly Met Pro Val Ser Gly Asp Arg Trp Pro Thr Gly Leu Arg  
 180 185 190

Ile Cys Leu Ala Glu Gln Leu Leu Glu Ser Leu Arg Ala Ala Gly Phe  
 195 200 205

Glu Ile Leu Asp Trp Glu Asp Val Ser Ser Arg Thr Arg Tyr Phe Met  
 210 215 220

Pro Gln Phe Ala Glu Glu Leu Ala Ala His Gln His Gly Ile Ala Asp  
 225 230 235 240

Arg Tyr Gly Pro Ala Val Ala Gly Trp Ala Ala Ala Val Cys Asp Tyr  
 245 250 255

Glu Lys Tyr Ala His Asp Met Gly Tyr Ala Ile Leu Thr Ala Arg Lys  
 260 265 270

Pro Val Gly  
 275

<210> 9  
 <211> 1173  
 <212> DNA  
 <213> Saccharopolyspora spinosa

<220>  
 <221> CDS  
 <222> (1)..(1170)  
 <223> ORF2; Glycosyltransferase

<400> 9  
 atg cgc gta ctc gtc gtt ccc ttg ccc tat ccg acg cat ctc atg gca 48  
 Met Arg Val Leu Val Val Pro Leu Pro Tyr Pro Thr His Leu Met Ala  
 1 5 10 15

atg gtg ccg ctg tgc tgg gcg ctg caa gca tcc ggg cac gag gtc ctg 96  
 Met Val Pro Leu Cys Trp Ala Leu Gln Ala Ser Gly His Glu Val Leu  
 20 25 30

atc gcc gca cca cca gag ctg cag gcg acc gcg cat ggt gca ggt ctc 144  
 Ile Ala Ala Pro Pro Glu Leu Gln Ala Thr Ala His Gly Ala Gly Leu  
 35 40 45

acc acg gcc ggg atc cgc ggg aac gac agg acc ggc gat acg ggt gga 192

Thr Thr Ala Gly Ile Arg Gly Asn Asp Arg Thr Gly Asn Thr Gly Gly  
 50 55 60

acc acg cag ctg cgc ttt ccc aat ccg gcg ttc ggt cag cgc gac acc 240  
 Thr Thr Gln Leu Arg Phe Pro Asn Pro Ala Phe Gly Gln Arg Asp Thr  
 65 70 75 80

gag gca ggc cgg caa ctg tgg gag cag acc gcg tcc aat gtc gcg caa 288  
 Glu Ala Gly Arg Gln Leu Trp Glu Gln Thr Ala Ser Asn Val Ala Gln  
 85 90 95

agc tcg ctc gat cag ctc ccc gaa tac ctt cga ctg gcc gag gcc tgg 336  
 Ser Ser Leu Asp Gln Leu Pro Glu Tyr Leu Arg Leu Ala Glu Ala Trp  
 100 105 110

cga ccg tca gtg ctg ttg gtc gac gtc tgc gcg ctg atc ggc cgg gtg 384  
 Arg Pro Ser Val Leu Leu Val Asp Val Cys Ala Leu Ile Gly Arg Val  
 115 120 125

ctc ggc gga ttg ctc gac ctg ccg gtc gtg ctg cac cgc tgg gga gtc 432  
 Leu Gly Gly Leu Leu Asp Leu Pro Val Val Leu His Arg Trp Gly Val  
 130 135 140

gac ccc acc gca ggc ccc ttc agc gat cga gcc cac gag ttg ctt gac 480  
 Asp Pro Thr Ala Gly Pro Phe Ser Asp Arg Ala His Glu Leu Leu Asp  
 145 150 155 160

ccg gtg tgc cgg cac cac gga ctg acc ggc ctg ccc act ccc gag ctc 528  
 Pro Val Cys Arg His His Gly Leu Thr Gly Leu Pro Thr Pro Glu Leu  
 165 170 175

atc ctc gat ccc tgt ccg ccg agc ctg caa gca agc gac gcg ccg caa 576  
 Ile Leu Asp Pro Cys Pro Pro Ser Leu Gln Ala Ser Asp Ala Pro Gln  
 180 185 190

ggc gca ccg gtc cag tac gtg ccg tac aac gga agc ggc gca ttc ccg 624  
 Gly Ala Pro Val Gln Tyr Val Pro Tyr Asn Gly Ser Gly Ala Phe Pro  
 195 200 205

gca tgg ggc gcg gcg cgc acc tca gca cgg cgg gtc tgc atc tgc atg 672  
 Ala Trp Gly Ala Ala Arg Thr Ser Ala Arg Arg Val Cys Ile Cys Met  
 210 215 220

ggc cgc atg gtg ctg aac gcc acc ggg ccg gct ccg ctg ctg cgc gca 720  
 Gly Arg Met Val Leu Asn Ala Thr Gly Pro Ala Pro Leu Leu Arg Ala  
 225 230 235 240

gta gcg gct gcc acc gag ttg ccc ggc gtc gag gcc gtg atc gcc gtt 768



Val Ala Ala Ala Thr Glu Leu Pro Gly Val Glu Ala Val Ile Ala Val  
 245 250 255

ccc ccc gag cac cgg gca ctt ctc acc gac cta ccc gac aac gcc cgg 816  
 Pro Pro Glu His Arg Ala Leu Leu Thr Asp Leu Pro Asp Asn Ala Arg  
 260 265 270

atc gcc gaa tcg gtc ccg ctc aac ctg ttc ctg cgt acc tgc gag ctg 864  
 Ile Ala Glu Ser Val Pro Leu Asn Leu Phe Leu Arg Thr Cys Glu Leu  
 275 280 285

gtc atc tgc gcg ggc ggc tcg gga acg gca ttc acc gcg acc cga ctc 912  
 Val Ile Cys Ala Gly Gly Ser Gly Thr Ala Phe Thr Ala Thr Arg Leu  
 290 295 300

ggc atc ccg caa ctc gtg ctt ccc cag tac ttc gac cag ttc gac tac 960  
 Gly Ile Pro Gln Leu Val Leu Pro Gln Tyr Phe Asp Gln Phe Asp Tyr  
 305 310 315 320

gcg cgc aac ctc gcc gct gcc ggg gcg ggc atc tgc ttg ccg gat gag 1008  
 Ala Arg Asn Leu Ala Ala Ala Gly Ala Gly Ile Cys Leu Pro Asp Glu  
 325 330 335

cag gcc cag tcc gac cac gaa cag ttc acc gac tca atc gca acg gtg 1056  
 Gln Ala Gln Ser Asp His Glu Gln Phe Thr Asp Ser Ile Ala Thr Val  
 340 345 350

ctc ggc gac acc ggc ttc gct gct gcg gca atc aaa ctc agc gac gag 1104  
 Leu Gly Asp Thr Gly Phe Ala Ala Ala Ala Ile Lys Leu Ser Asp Glu  
 355 360 365

atc acg gcc atg ccc cat ccc gcc gcg ctg gtg cgg acg ctg gag aac 1152  
 Ile Thr Ala Met Pro His Pro Ala Ala Leu Val Arg Thr Leu Glu Asn  
 370 375 380

act gcg gcc atc cgt gcc tga 1173  
 Thr Ala Ala Ile Arg Ala  
 385 390

<210> 10

<211> 390

<212> PRT

<213> Saccharopolyspora spinosa

<400> 10

Met Arg Val Leu Val Val Pro Leu Pro Tyr Pro Thr His Leu Met Ala  
 1 5 10 15

Met Val Pro Leu Cys Trp Ala Leu Gln Ala Ser Gly His Glu Val Leu  
20 25 30

Ile Ala Ala Pro Pro Glu Leu Gln Ala Thr Ala His Gly Ala Gly Leu  
35 40 45

Thr Thr Ala Gly Ile Arg Gly Asn Asp Arg Thr Gly Asp Thr Gly Gly  
50 55 60

Thr Thr Gln Leu Arg Phe Pro Asn Pro Ala Phe Gly Gln Arg Asp Thr  
65 70 75 80

Glu Ala Gly Arg Gln Leu Trp Glu Gln Thr Ala Ser Asn Val Ala Gln  
85 90 95

Ser Ser Leu Asp Gln Leu Pro Glu Tyr Leu Arg Leu Ala Glu Ala Trp  
100 105 110

Arg Pro Ser Val Leu Leu Val Asp Val Cys Ala Leu Ile Gly Arg Val  
115 120 125

Leu Gly Gly Leu Leu Asp Leu Pro Val Val Leu His Arg Trp Gly Val  
130 135 140

Asp Pro Thr Ala Gly Pro Phe Ser Asp Arg Ala His Glu Leu Leu Asp  
145 150 155 160

Pro Val Cys Arg His His Gly Leu Thr Gly Leu Pro Thr Pro Glu Leu  
165 170 175

Ile Leu Asp Pro Cys Pro Pro Ser Leu Gln Ala Ser Asp Ala Pro Gln  
180 185 190

Gly Ala Pro Val Gln Tyr Val Pro Tyr Asn Gly Ser Gly Ala Phe Pro  
195 200 205

Ala Trp Gly Ala Ala Arg Thr Ser Ala Arg Arg Val Cys Ile Cys Met  
210 215 220

Gly Arg Met Val Leu Asn Ala Thr Gly Pro Ala Pro Leu Leu Arg Ala  
225 230 235 240

Val Ala Ala Ala Thr Glu Leu Pro Gly Val Glu Ala Val Ile Ala Val  
245 250 255

Pro Pro Glu His Arg Ala Leu Leu Thr Asp Leu Pro Asp Asn Ala Arg  
260 265 270

Ile Ala Glu Ser Val Pro Leu Asn Leu Phe Leu Arg Thr Cys Glu Leu  
275 280 285

Val Ile Cys Ala Gly Gly Ser Gly Thr Ala Phe Thr Ala Thr Arg Leu  
290 295 300

Gly Ile Pro Gln Leu Val Leu Pro Gln Tyr Phe Asp Gln Phe Asp Tyr  
305 310 315 320

Ala Arg Asn Leu Ala Ala Ala Gly Ala Gly Ile Cys Leu Pro Asp Glu  
325 330 335

Gln Ala Gln Ser Asp His Glu Gln Phe Thr Asp Ser Ile Ala Thr Val  
340 345 350

Leu Gly Asp Thr Gly Phe Ala Ala Ala Ala Ile Lys Leu Ser Asp Glu  
355 360 365

Ile Thr Ala Met Pro His Pro Ala Ala Leu Val Arg Thr Leu Glu Asn  
370 375 380

Thr Ala Ala Ile Arg Ala  
385 390

<210> 11  
<211> 753  
<212> DNA  
<213> Saccharopolyspora spinosa

<220>  
<221> CDS  
<222> (1)..(750)  
<223> ORF3; O-Methyltransferase

<400> 11  
atg ccc tcc cag aac gcg ctg tac ctg gac ctg ctc aag aag gta ctc 48  
Met Pro Ser Gln Asn Ala Leu Tyr Leu Asp Leu Leu Lys Lys Val Leu  
1 5 10 15

---

acc aac acg att tac agt gat cgg ccg cat ccg aac gcc tgg cag gac 96  
Thr Asn Thr Ile Tyr Ser Asp Arg Pro His Pro Asn Ala Trp Gln Asp  
20 25 30

aac acc gac tac agg cag gcc gct cgg gcc aaa ggc acg gac tgg cca 144  
Asn Thr Asp Tyr Arg Gln Ala Ala Arg Ala Lys Gly Thr Asp Trp Pro

act gtc gcg cac acg atg atc ggt ctg gag cgg ctg gac aac ctc cag 192  
 Thr Val Ala His Thr Met Ile Gly Leu Glu Arg Leu Asp Asn Leu Gln  
 50 55 60

cac tgc gtg gaa gcc gtg ctc gca gac ggt gtt ccc ggg gat ttc gcc 240  
 His Cys Val Glu Ala Val Leu Ala Asp Gly Val Pro Gly Asp Phe Ala  
 65 70 75 80

gag acc ggt gtc tgg cgg ggc ggc gca tgc atc ttc atg cgc gcg gtt 288  
 Glu Thr Gly Val Trp Arg Gly Gly Ala Cys Ile Phe Met Arg Ala Val  
 85 90 95

ctc cag gca ttc gga gat acc gga cgt acc gtc tgg gta gtg gat tcc 336  
 Leu Gln Ala Phe Gly Asp Thr Gly Arg Thr Val Trp Val Val Asp Ser  
 100 105 110

ttc cag gga atg ccg gaa agc tct gcg caa gac cac caa gcg gac cag 384  
 Phe Gln Gly Met Pro Glu Ser Ser Ala Gln Asp His Gln Ala Asp Gln  
 115 120 125

gct atg gcg ctg cac gag tac aac gac gtg ctt ggc gta tcg ctt gag 432  
 Ala Met Ala Leu His Glu Tyr Asn Asp Val Leu Gly Val Ser Leu Glu  
 130 135 140

acc gtc cgg cag aac ttc gcc cgc tac ggg ctg ctc gac gaa cag gtc 480  
 Thr Val Arg Gln Asn Phe Ala Arg Tyr Gly Leu Leu Asp Glu Gln Val  
 145 150 155 160

agg ttc ctc ccc ggc tgg ttc cgg gac acc ttg ccc acc gcc ccc atc 528  
 Arg Phe Leu Pro Gly Trp Phe Arg Asp Thr Leu Pro Thr Ala Pro Ile  
 165 170 175

cag gaa ctc gcc gtg cta cga ctc gac ggc gac ctc tac gaa tcc aca 576  
 Gln Glu Leu Ala Val Leu Arg Leu Asp Gly Asp Leu Tyr Glu Ser Thr  
 180 185 190

atg gac tca ttg cgg aac ctg tac ccg aag ctc tcg ccg ggc gga ttc 624  
 Met Asp Ser Leu Arg Asn Leu Tyr Pro Lys Leu Ser Pro Gly Gly Phe  
 195 200 205

gtc atc atc gac gac tat ttt ctg ccg tcc tgc cag gac gcg gtg aag 672  
 Val Ile Ile Asp Asp Tyr Phe Leu Pro Ser Cys Gln Asp Ala Val Lys  
 210 215 220

ggg ttc cgc gcg gaa ctc ggg atc acg gaa ccc atc cac gac atc gac 720  
 Gly Phe Arg Ala Glu Leu Gly Ile Thr Glu Pro Ile His Asp Ile Asp

225

230

235

240

ggc acg ggc gcc tac tgg cgc cgc agc tgg tga  
 Gly Thr Gly Ala Tyr Trp Arg Arg Ser Trp  
 245 250

753

&lt;210&gt; 12

&lt;211&gt; 250

&lt;212&gt; PRT

&lt;213&gt; Saccharopolyspora spinosa

&lt;400&gt; 12

Met Pro Ser Gln Asn Ala Leu Tyr Leu Asp Leu Leu Lys Lys Val Leu  
 1 5 10 15

Thr Asn Thr Ile Tyr Ser Asp Arg Pro His Pro Asn Ala Trp Gln Asp  
 20 25 30

Asn Thr Asp Tyr Arg Gln Ala Ala Arg Ala Lys Gly Thr Asp Trp Pro  
 35 40 45

Thr Val Ala His Thr Met Ile Gly Leu Glu Arg Leu Asp Asn Leu Gln  
 50 55 60

His Cys Val Glu Ala Val Leu Ala Asp Gly Val Pro Gly Asp Phe Ala  
 65 70 75 80

Glu Thr Gly Val Trp Arg Gly Gly Ala Cys Ile Phe Met Arg Ala Val  
 85 90 95

Leu Gln Ala Phe Gly Asp Thr Gly Arg Thr Val Trp Val Val Asp Ser  
 100 105 110

Phe Gln Gly Met Pro Glu Ser Ser Ala Gln Asp His Gln Ala Asp Gln  
 115 120 125

Ala Met Ala Leu His Glu Tyr Asn Asp Val Leu Gly Val Ser Leu Glu  
 130 135 140

Thr Val Arg Gln Asn Phe Ala Arg Tyr Gly Leu Leu Asp Glu Gln Val  
 145 150 155 160

Arg Phe Leu Pro Gly Trp Phe Arg Asp Thr Leu Pro Thr Ala Pro Ile  
 165 170 175

Gln Glu Leu Ala Val Leu Arg Leu Asp Gly Asp Leu Tyr Glu Ser Thr  
 180 185 190

Met Asp Ser Leu Arg Asn Leu Tyr Pro Lys Leu Ser Pro Gly Gly Phe  
 195 200 205

Val Ile Ile Asp Asp Tyr Phe Leu Pro Ser Cys Gln Asp Ala Val Lys  
 210 215 220

Gly Phe Arg Ala Glu Leu Gly Ile Thr Glu Pro Ile His Asp Ile Asp  
 225 230 235 240

Gly Thr Gly Ala Tyr Trp Arg Arg Ser Trp  
 245 250

<210> 13

<211> 1188

<212> DNA

<213> Saccharopolyspora spinosa

<220>

<221> CDS

<222> (1)..(1185)

<223> ORF4; O-Methyltransferase

<400> 13

atg agt gag atc gca gtt gcc ccc tgg tcg gtg gtg gag cgt ttg ctg 48  
 Met Ser Glu Ile Ala Val Ala Pro Trp Ser Val Val Glu Arg Leu Leu  
 1 5 10 15

ctc gcg gcg ggt gcg ggc ccg gcg aag ctc cag gaa gca gtg cag gtg 96  
 Leu Ala Ala Gly Ala Gly Pro Ala Lys Leu Gln Glu Ala Val Gln Val  
 20 25 30

gcc gga ctg gac gcg gtg gcc gac gcc atc gtc gac gaa ctc gtc gta 144  
 Ala Gly Leu Asp Ala Val Ala Asp Ala Ile Val Asp Glu Leu Val Val  
 35 40 45

cgc tgc gat ccg ctg tcg ttg gac gag tcg gtg cga atc ggc ctg gag 192  
 Arg Cys Asp Pro Leu Ser Leu Asp Glu Ser Val Arg Ile Gly Leu Glu  
 50 55 60

---

atc act tct ggc gct cag ctg gtc cgg aga acc gtt gag ctc gat cac 240  
 Ile Thr Ser Gly Ala Gln Leu Val Arg Arg Thr Val Glu Leu Asp His  
 65 70 75 80

gca ggc ctg cgg ctc gcg gcg gtc gcc gaa gca gct gct gtt ctc cgg 288  
 Ala Gly Leu Arg Leu Ala Ala Val Ala Glu Ala Ala Val Leu Arg

ttc gac gcg gtg gat ctg ctg gaa ggg ctc ttc ggc ccg gtt gac ggc 336  
 Phe Asp Ala Val Asp Leu Leu Glu Gly Leu Phe Gly Pro Val Asp Gly  
 100 105 110

agg cgg cac aac agc cgt gaa gtc cgc tgg tcg gac agc atg acg cag 384  
 Arg Arg His Asn Ser Arg Glu Val Arg Trp Ser Asp Ser Met Thr Gln  
 115 120 125

ttc tcg ccc gac cag ggc ctc gcc ggc gcg cag cgc ctg ctg gcg ttc 432  
 Phe Ser Pro Asp Gln Gly Leu Ala Gly Ala Gln Arg Leu Leu Ala Phe  
 130 135 140

cgg aac agg gtg tcc acc gcg gtg cac gcc gtg ctg gcc gca gcc gcc 480  
 Arg Asn Arg Val Ser Thr Ala Val His Ala Val Leu Ala Ala Ala Ala  
 145 150 155 160

acc agg cgc gcg gac ctc ggt gcg ctg gca gtc cgc tac gga tcc gac 528  
 Thr Arg Arg Ala Asp Leu Gly Ala Leu Ala Val Arg Tyr Gly Ser Asp  
 165 170 175

aaa tgg gcg gac ctg cac tgg tac acc gaa cac tac gag cac cac ttc 576  
 Lys Trp Ala Asp Leu His Trp Tyr Thr Glu His Tyr Glu His His Phe  
 180 185 190

tcc cga ttc cag gat gcc ccg gtg cga gtg ttg gaa ata gga atc ggt 624  
 Ser Arg Phe Gln Asp Ala Pro Val Arg Val Leu Glu Ile Gly Ile Gly  
 195 200 205

ggt tat cac gca ccc gaa ctc ggt ggt gct tcg ctg cgc atg tgg cag 672  
 Gly Tyr His Ala Pro Glu Leu Gly Gly Ala Ser Leu Arg Met Trp Gln  
 210 215 220

cgg tac ttc cgg cga ggt ctc gtt tac ggg ctg gac att ttc gag aaa 720  
 Arg Tyr Phe Arg Arg Gly Leu Val Tyr Gly Leu Asp Ile Phe Glu Lys  
 225 230 235 240

gcc ggg aac gaa ggg cac cga gtg cga aag ctg cga ggt gac cag agc 768  
 Ala Gly Asn Glu Gly His Arg Val Arg Lys Leu Arg Gly Asp Gln Ser  
 245 250 255

gat gcg gaa ttc ctg gaa gac atg gcg ggg aag atc ggg ccg ttc gac 816  
 Asp Ala Glu Phe Leu Glu Asp Met Ala Gly Lys Ile Gly Pro Phe Asp  
 260 265 270

att gtc atc gac gac ggc agc cat gtc aac gac cac gtc aag aaa tcc 864  
 Ile Val Ile Asp Asp Gly Ser His Val Asn Asp His Val Lys Lys Ser

275

280

285

ttc caa tcc ctg ttt ccg cac gtc cgc cca ggt ggt ttg tac gtc atc 912  
 Phe Gln Ser Leu Phe Pro His Val Arg Pro Gly Gly Leu Tyr Val Ile  
 290 295 300

gag gat ctc cag acg gcg tac tgg ccc ggc tac ggc ggt cgc gat ggg 960  
 Glu Asp Leu Gln Thr Ala Tyr Trp Pro Gly Tyr Gly Gly Arg Asp Gly  
 305 310 315 320

gaa ccc gcg gcc cag cgc acc tcg atc gac atg ctc aaa gaa ctg atc 1008  
 Glu Pro Ala Ala Gln Arg Thr Ser Ile Asp Met Leu Lys Glu Leu Ile  
 325 330 335

gac ggc ctg cat tat cag gag cgc gaa tcg cgg tgc ggg acc gag ccc 1056  
 Asp Gly Leu His Tyr Gln Glu Arg Glu Ser Arg Cys Gly Thr Glu Pro  
 340 345 350

tcc tac acg gaa cgg aac gtg gcg gcc ctg cac ttc tac cac aac ctg 1104  
 Ser Tyr Thr Glu Arg Asn Val Ala Ala Leu His Phe Tyr His Asn Leu  
 355 360 365

gta ttc gtg gag aaa ggg ctc aac gct gag cct gcc gcg ccg ggg ttc 1152  
 Val Phe Val Glu Lys Gly Leu Asn Ala Glu Pro Ala Ala Pro Gly Phe  
 370 375 380

gtg ccc cgg caa gcg ctc ggc gtc gag gac ggc tga 1188  
 Val Pro Arg Gln Ala Leu Gly Val Glu Asp Gly  
 385 390 395

&lt;210&gt; 14

&lt;211&gt; 395

&lt;212&gt; PRT

&lt;213&gt; Saccharopolyspora spinosa

&lt;400&gt; 14

Met Ser Glu Ile Ala Val Ala Pro Trp Ser Val Val Glu Arg Leu Leu  
 1 5 10 15

Leu Ala Ala Gly Ala Gly Pro Ala Lys Leu Gln Glu Ala Val Gln Val  
 20 25 30

Ala Gly Leu Asp Ala Val Ala Asp Ala Ile Val Asp Glu Leu Val Val  
 35 40 45

Arg Cys Asp Pro Leu Ser Leu Asp Glu Ser Val Arg Ile Gly Leu Glu  
 50 55 60



Ile Thr Ser Gly Ala Gln Leu Val Arg Arg Thr Val Glu Leu Asp His  
65 70 75 80

Ala Gly Leu Arg Leu Ala Ala Val Ala Glu Ala Ala Val Leu Arg  
85 90 95

Phe Asp Ala Val Asp Leu Leu Glu Gly Leu Phe Gly Pro Val Asp Gly  
100 105 110

Arg Arg His Asn Ser Arg Glu Val Arg Trp Ser Asp Ser Met Thr Gln  
115 120 125

Phe Ser Pro Asp Gln Gly Leu Ala Gly Ala Gln Arg Leu Leu Ala Phe  
130 135 140

Arg Asn Arg Val Ser Thr Ala Val His Ala Val Leu Ala Ala Ala Ala  
145 150 155 160

Thr Arg Arg Ala Asp Leu Gly Ala Leu Ala Val Arg Tyr Gly Ser Asp  
165 170 175

Lys Trp Ala Asp Leu His Trp Tyr Thr Glu His Tyr Glu His His Phe  
180 185 190

Ser Arg Phe Gln Asp Ala Pro Val Arg Val Leu Glu Ile Gly Ile Gly  
195 200 205

Gly Tyr His Ala Pro Glu Leu Gly Gly Ala Ser Leu Arg Met Trp Gln  
210 215 220

Arg Tyr Phe Arg Arg Gly Leu Val Tyr Gly Leu Asp Ile Phe Glu Lys  
225 230 235 240

Ala Gly Asn Glu Gly His Arg Val Arg Lys Leu Arg Gly Asp Gln Ser  
245 250 255

Asp Ala Glu Phe Leu Glu Asp Met Ala Gly Lys Ile Gly Pro Phe Asp  
260 265 270

Ile Val Ile Asp Asp Gly Ser His Val Asn Asp His Val Lys Lys Ser  
275 280 285

Phe Gln Ser Leu Phe Pro His Val Arg Pro Gly Gly Leu Tyr Val Ile  
290 295 300

Glu Asp Leu Gln Thr Ala Tyr Trp Pro Gly Tyr Gly Gly Arg Asp Gly  
305 310 315 320

Glu Pro Ala Ala Gln Arg Thr Ser Ile Asp Met Leu Lys Glu Leu Ile  
 325 330 335

Asp Gly Leu His Tyr Gln Glu Arg Glu Ser Arg Cys Gly Thr Glu Pro  
 340 345 350

Ser Tyr Thr Glu Arg Asn Val Ala Ala Leu His Phe Tyr His Asn Leu  
 355 360 365

Val Phe Val Glu Lys Gly Leu Asn Ala Glu Pro Ala Ala Pro Gly Phe  
 370 375 380

Val Pro Arg Gln Ala Leu Gly Val Glu Asp Gly  
 385 390 395

<210> 15

<211> 1620

<212> DNA

<213> Saccharopolyspora spinosa

<220>

<221> CDS

<222> (1)..(1617)

<223> ORF5; C-C verknüpfendes Enzym, Cyclisierungsenzym

<400> 15

atg atc tcg gct gcg ggc gaa caa agt gga cca gtc aga aaa gga ggg 48  
 Met Ile Ser Ala Ala Gly Glu Gln Ser Gly Pro Val Arg Lys Gly Gly  
 1 5 10 15

gcg gtg ccc gaa ttc cat gac ccg gca ccc atg aat cgt cga acc cca 96  
 Ala Val Pro Glu Phe His Asp Pro Ala Pro Met Asn Arg Arg Thr Pro  
 20 25 30

gga aca gag atc acc gtc gag ccc gac gat cct cgt tat ccg gac ctc 144  
 Gly Thr Glu Ile Thr Val Glu Pro Asp Asp Pro Arg Tyr Pro Asp Leu  
 35 40 45

~~gtc gtc ggg cac aac ccc cgt ttc acc gga aaa ccc gaa cgc atc cac 192~~  
 Val Val Gly His Asn Pro Arg Phe Thr Gly Lys Pro Glu Arg Ile His  
 50 55 60

atc gcc agc tcc gcc gaa gac gtc gtg cac gcc gtc gcc gac gcc gtg 240  
 Ile Ala Ser Ser Ala Glu Asp Val Val His Ala Val Ala Asp Ala Val  
 65 70 75 80

cgc acc ggc agg cgg gta ggg gtc cgc agc ggc ggg cac tgc ttc gag 288  
 Arg Thr Gly Arg Arg Val Gly Val Arg Ser Gly Gly His Cys Phe Glu  
                     85                    90                    95

aat ctc gtt gcg gac ccg gcg atc cga gtg ctc gtc gac ctc tcc gag 336  
 Asn Leu Val Ala Asp Pro Ala Ile Arg Val Leu Val Asp Leu Ser Glu  
                     100                    105                    110

ctc aac cgc gtg tac tac gac agc acg cgc ggg gca ttc gcg atc gag 384  
 Leu Asn Arg Val Tyr Tyr Asp Ser Thr Arg Gly Ala Phe Ala Ile Glu  
                     115                    120                    125

gcg ggc gcc gcc ctc ggg cag gtg tac cga acc ctg ttc aag aac tgg 432  
 Ala Gly Ala Ala Leu Gly Gln Val Tyr Arg Thr Leu Phe Lys Asn Trp  
                     130                    135                    140

ggc gtg acg atc ccg acc ggc gca tgt ccc ggg gtg ggc gca ggc ggg 480  
 Gly Val Thr Ile Pro Thr Gly Ala Cys Pro Gly Val Gly Ala Gly Gly  
                     145                    150                    155                    160

cac atc ctc ggc ggg gga tac ggc ccg ctg tcg cgc cga ttc ggt tcg 528  
 His Ile Leu Gly Gly Gly Tyr Gly Pro Leu Ser Arg Arg Phe Gly Ser  
                     165                    170                    175

gtc gtc gac tac ctt caa ggc gtc gag gtc gtc gtg gtc gac cag gcc 576  
 Val Val Asp Tyr Leu Gln Gly Val Glu Val Val Val Val Asp Gln Ala  
                     180                    185                    190

ggt gaa gtg cac atc gtc gag gcc gac ccg aac tcc acg ggc gcc ggt 624  
 Gly Glu Val His Ile Val Glu Ala Asp Arg Asn Ser Thr Gly Ala Gly  
                     195                    200                    205

cac gac ttg tgg tgg gcg cac acc ggt ggc ggt ggc ggc aac ttc ggg 672  
 His Asp Leu Trp Trp Ala His Thr Gly Gly Gly Gly Gly Asn Phe Gly  
                     210                    215                    220

atc gtc acc agg ttt tgg ctc cga acg ccg gac gtg gtc agc acc gac 720  
 Ile Val Thr Arg Phe Trp Leu Arg Thr Pro Asp Val Val Ser Thr Asp  
                     225                    230                    235                    240

~~gcc gca gag ctc ctg cca cgg ccg ccc gcg aca gtg ctg ctc cga tcg 768~~  
~~Ala Ala Glu Leu Leu Pro Arg Pro Pro Ala Thr Val Leu Leu Arg Ser~~  
~~245                    250                    255~~

ttc cac tgg ccg tgg cac gaa ctg aca gag cag tca ttc gcc gtg ctc 816  
 Phe His Trp Pro Trp His Glu Leu Thr Glu Gln Ser Phe Ala Val Leu  
                     260                    265                    270

cta cag aac ttc ggc aat tgg tac gag cag cac agc gcg cct gaa tcc	864
Leu Gln Asn Phe Gly Asn Trp Tyr Glu Gln His Ser Ala Pro Glu Ser	
275 280 285	
acg caa ctc ggg ttg ttc agc acg ctc gtc tgc gca cac cgg caa gct	912
Thr Gln Leu Gly Leu Phe Ser Thr Leu Val Cys Ala His Arg Gln Ala	
290 295 300	
ggc tac gtc acg ctg aac gtt cac ctg gac ggc acg gat ccg aac gcg	960
Gly Tyr Val Thr Leu Asn Val His Leu Asp Gly Thr Asp Pro Asn Ala	
305 310 315 320	
gaa cgc acc ctg gcc gaa cac ctg tcg gcg atc aac gcc cag gtc ggc	1008
Glu Arg Thr Leu Ala Glu His Leu Ser Ala Ile Asn Ala Gln Val Gly	
325 330 335	
gtg act cca gcc gaa ggg ctg cgg gaa acc ctg ccg tgg ttg cga tcg	1056
Val Thr Pro Ala Glu Gly Leu Arg Glu Thr Leu Pro Trp Leu Arg Ser	
340 345 350	
acc cag gtg gcc ggg gcg atc gcc gaa ggc ggc gaa ccg ggc atg caa	1104
Thr Gln Val Ala Gly Ala Ile Ala Glu Gly Gly Glu Pro Gly Met Gln	
355 360 365	
cgg acc aag gtc aaa gcc gcc tac ttg cgc acc ggg ctg tcc gaa gct	1152
Arg Thr Lys Val Lys Ala Ala Tyr Leu Arg Thr Gly Leu Ser Glu Ala	
370 375 380	
caa cta gcc acg gtt tac cgg cgg ctg acc gtc tac gga tac gac aac	1200
Gln Leu Ala Thr Val Tyr Arg Arg Leu Thr Val Tyr Gly Tyr Asp Asn	
385 390 395 400	
cct gcg gcg gcg ctg ttg ctg ctc ggt tac ggc ggt atg gcg aat gcc	1248
Pro Ala Ala Ala Leu Leu Leu Gly Tyr Gly Gly Met Ala Asn Ala	
405 410 415	
gtg gct ccg tcg gcc acc gca ctc gct cag cgc gac tcg gtt ctc aaa	1296
Val Ala Pro Ser Ala Thr Ala Leu Ala Gln Arg Asp Ser Val Leu Lys	
420 425 430	
gcg ctg ttc gtc acg aac tgg tcg gag ccc gcc gag gac gag cgg cat	1344
Ala Leu Phe Val Thr Asn Trp Ser Glu Pro Ala Glu Asp Glu Arg His	
435 440 445	
ctg acc tgg att cgc ggt ttc tac cgc gag atg tac gcc gaa acc ggc	1392
Leu Thr Trp Ile Arg Gly Phe Tyr Arg Glu Met Tyr Ala Glu Thr Gly	
450 455 460	

gga gtt ccg gtg cca ggt acc cgt gtc gac ggc tcc tac atc aac tac 1440  
 Gly Val Pro Val Pro Gly Thr Arg Val Asp Gly Ser Tyr Ile Asn Tyr  
 465 470 475 480

ccg gac acc gac ttg gcc gat cca ttg tgg aac acc tcc ggt gtt gcc 1488  
 Pro Asp Thr Asp Leu Ala Asp Pro Leu Trp Asn Thr Ser Gly Val Ala  
 485 490 495

tgg cac gac ctg tac tac aaa gac aac tac ccg cgg ctg cag cgg gcc 1536  
 Trp His Asp Leu Tyr Tyr Lys Asp Asn Tyr Pro Arg Leu Gln Arg Ala  
 500 505 510

aaa gcg cgg tgg gat ccg cag aac atc ttc cag cac ggc ctg tcg atc 1584  
 Lys Ala Arg Trp Asp Pro Gln Asn Ile Phe Gln His Gly Leu Ser Ile  
 515 520 525

aaa ccg ccg gca cgg ctt tca ccc ggt cag cca tga 1620  
 Lys Pro Pro Ala Arg Leu Ser Pro Gly Gln Pro  
 530 535

<210> 16  
 <211> 539  
 <212> PRT  
 <213> Saccharopolyspora spinosa

<400> 16  
 Met Ile Ser Ala Ala Gly Glu Gln Ser Gly Pro Val Arg Lys Gly Gly  
 1 5 10 15

Ala Val Pro Glu Phe His Asp Pro Ala Pro Met Asn Arg Arg Thr Pro  
 20 25 30

Gly Thr Glu Ile Thr Val Glu Pro Asp Asp Pro Arg Tyr Pro Asp Leu  
 35 40 45

Val Val Gly His Asn Pro Arg Phe Thr Gly Lys Pro Glu Arg Ile His  
 50 55 60

Ile Ala Ser Ser Ala Glu Asp Val Val His Ala Val Ala Asp Ala Val  
 65 70 75 80

Arg Thr Gly Arg Arg Val Gly Val Arg Ser Gly Gly His Cys Phe Glu  
 85 90 95

Asn Leu Val Ala Asp Pro Ala Ile Arg Val Leu Val Asp Leu Ser Glu  
 100 105 110

Leu Asn Arg Val Tyr Tyr Asp Ser Thr Arg Gly Ala Phe Ala Ile Glu  
115 120 125

Ala Gly Ala Ala Leu Gly Gln Val Tyr Arg Thr Leu Phe Lys Asn Trp  
130 135 140

Gly Val Thr Ile Pro Thr Gly Ala Cys Pro Gly Val Gly Ala Gly Gly  
145 150 155 160

His Ile Leu Gly Gly Gly Tyr Gly Pro Leu Ser Arg Arg Phe Gly Ser  
165 170 175

Val Val Asp Tyr Leu Gln Gly Val Glu Val Val Val Val Asp Gln Ala  
180 185 190

Gly Glu Val His Ile Val Glu Ala Asp Arg Asn Ser Thr Gly Ala Gly  
195 200 205

His Asp Leu Trp Trp Ala His Thr Gly Gly Gly Gly Gly Asn Phe Gly  
210 215 220

Ile Val Thr Arg Phe Trp Leu Arg Thr Pro Asp Val Val Ser Thr Asp  
225 230 235 240

Ala Ala Glu Leu Leu Pro Arg Pro Pro Ala Thr Val Leu Leu Arg Ser  
245 250 255

Phe His Trp Pro Trp His Glu Leu Thr Glu Gln Ser Phe Ala Val Leu  
260 265 270

Leu Gln Asn Phe Gly Asn Trp Tyr Glu Gln His Ser Ala Pro Glu Ser  
275 280 285

Thr Gln Leu Gly Leu Phe Ser Thr Leu Val Cys Ala His Arg Gln Ala  
290 295 300

Gly Tyr Val Thr Leu Asn Val His Leu Asp Gly Thr Asp Pro Asn Ala  
305 310 315 320

Glu Arg Thr Leu Ala Glu His Leu Ser Ala Ile Asn Ala Gln Val Gly  
325 330 335

Val Thr Pro Ala Glu Gly Leu Arg Glu Thr Leu Pro Trp Leu Arg Ser  
340 345 350

Thr Gln Val Ala Gly Ala Ile Ala Glu Gly Gly Glu Pro Gly Met Gln  
355 360 365

Arg Thr Lys Val Lys Ala Ala Tyr Leu Arg Thr Gly Leu Ser Glu Ala  
370 375 380

Gln Leu Ala Thr Val Tyr Arg Arg Leu Thr Val Tyr Gly Tyr Asp Asn  
385 390 395 400

Pro Ala Ala Ala Leu Leu Leu Leu Gly Tyr Gly Gly Met Ala Asn Ala  
405 410 415

Val Ala Pro Ser Ala Thr Ala Leu Ala Gln Arg Asp Ser Val Leu Lys  
420 425 430

Ala Leu Phe Val Thr Asn Trp Ser Glu Pro Ala Glu Asp Glu Arg His  
435 440 445

Leu Thr Trp Ile Arg Gly Phe Tyr Arg Glu Met Tyr Ala Glu Thr Gly  
450 455 460

Gly Val Pro Val Pro Gly Thr Arg Val Asp Gly Ser Tyr Ile Asn Tyr  
465 470 475 480

Pro Asp Thr Asp Leu Ala Asp Pro Leu Trp Asn Thr Ser Gly Val Ala  
485 490 495

Trp His Asp Leu Tyr Tyr Lys Asp Asn Tyr Pro Arg Leu Gln Arg Ala  
500 505 510

Lys Ala Arg Trp Asp Pro Gln Asn Ile Phe Gln His Gly Leu Ser Ile  
515 520 525

Lys Pro Pro Ala Arg Leu Ser Pro Gly Gln Pro  
530 535

<210> 17  
<211> 1194  
<212> DNA  
<213> Saccharopolyspora spinosa

<220>

<221> CDS  
<222> (1)..(1191)  
<223> ORF6; Methyltransferase

<400> 17

atg tcc aca acg cac gag atc gaa acc gtg gaa cgc atc atc ctc gcc 48

Met	Ser	Thr	Thr	His	Glu	Ile	Glu	Thr	Val	Glu	Arg	Ile	Ile	Leu	Ala	
1				5					10					15		
gcc gga tcc agt gcg gcg agc ctg gcc gac ctg acc acc gaa ctc gga																96
Ala Gly Ser Ser Ala Ala Ser Leu Ala Asp Leu Thr Thr Glu Leu Gly																
				20					25					30		
ctc gcc agg atc gca ccc gtg ctg atc gac gag atc ctc ttc cgc gcg																144
Leu Ala Arg Ile Ala Pro Val Leu Ile Asp Glu Ile Leu Phe Arg Ala																
				35					40					45		
gaa ccg gcc ccc gac atc gaa cgg acc gag gtc gcg gtc cag atc acc																192
Glu Pro Ala Pro Asp Ile Glu Arg Thr Glu Val Ala Val Gln Ile Thr																
				50					55					60		
cac cga ggc gag acc gtt gac ttc gtc ctg acg cta cag tcc ggt gag																240
His Arg Gly Glu Thr Val Asp Phe Val Leu Thr Leu Gln Ser Gly Glu																
				65					70					75	80	
ctg atc aag gcc gag caa cga ccg gtc gga gac gtc ccg ctg cgg atc																288
Leu Ile Lys Ala Glu Gln Arg Pro Val Gly Asp Val Pro Leu Arg Ile																
				85					90					95		
ggt tac gag ctc acc gat ctc atc gcc gag ttg ttc ggc cca gga gct																336
Gly Tyr Glu Leu Thr Asp Leu Ile Ala Glu Leu Phe Gly Pro Gly Ala																
				100					105					110		
ccc agg gcc gtc ggc gcc cgg agc acc aac ttc ctc cga acc acc aca																384
Pro Arg Ala Val Gly Ala Arg Ser Thr Asn Phe Leu Arg Thr Thr Thr																
				115					120					125		
tcc ggt tcg ata ccc ggt ccg tcg gaa ctg tcc gat ggc ttc cag gcc																432
Ser Gly Ser Ile Pro Gly Pro Ser Glu Leu Ser Asp Gly Phe Gln Ala																
				130					135					140		
atc tcc gca gtg gtc gcc ggc tgc ggg cac cga cgt ccc gac ctc aac																480
Ile Ser Ala Val Val Ala Gly Cys Gly His Arg Arg Pro Asp Leu Asn																
				145					150					155	160	
ttg ctc gcc tcc cac tac cgc acg gac aag tgg ggc ggc ctg cac tgg																528
Leu Leu Ala Ser His Tyr Arg Thr Asp Lys Trp Gly Gly Leu His Trp																
				165					170					175		
ttc acc ccg cta tac gag cga cac ctc ggc gag ttc cgt gat cgc ccg																576
Phe Thr Pro Leu Tyr Glu Arg His Leu Gly Glu Phe Arg Asp Arg Pro																
				180					185					190		
gtg cgc atc ctg gag atc ggt gtc ggt ggc tac aac ttc gac ggt ggc																624



Val Arg Ile Leu Glu Ile Gly	Val Gly Gly Tyr Asn Phe Asp Gly Gly	
195	200	205
ggc ggc gaa tcc ctg aag atg tgg aag cgc tac ttc cac cgc ggc ctc	672	
Gly Gly Glu Ser Leu Lys Met Trp Lys Arg Tyr Phe His Arg Gly Leu		
210	215	220
gtg ttc ggg atg gac gtt ttc gac aag tcc ttc ctc gac cag cag agg	720	
Val Phe Gly Met Asp Val Phe Asp Lys Ser Phe Leu Asp Gln Gln Arg		
225	230	235
ctc tgc acc gtc cgc gcc gac cag agc aag ccc gag gag ctg gcc gcc	768	
Leu Cys Thr Val Arg Ala Asp Gln Ser Lys Pro Glu Glu Leu Ala Ala		
245	250	255
gtt gac gac aag tac gga ccg ttc gac atc atc atc gac gat ggc agc	816	
Val Asp Asp Lys Tyr Gly Pro Phe Asp Ile Ile Ile Asp Asp Gly Ser		
260	265	270
cac atc aac gga cac gtg cgc aca tcc ctg gaa acg ctg ttc ccc cgg	864	
His Ile Asn Gly His Val Arg Thr Ser Leu Glu Thr Leu Phe Pro Arg		
275	280	285
ttg cgc agc ggt ggc gta tac gtg atc gag gat ctg tgg acg acc tat	912	
Leu Arg Ser Gly Gly Val Tyr Val Ile Glu Asp Leu Trp Thr Thr Tyr		
290	295	300
gct ccc gga ttc ggc ggg cag gcg cag tgc ccg gcc gca ccc ggc acc	960	
Ala Pro Gly Phe Gly Gly Gln Ala Gln Cys Pro Ala Ala Pro Gly Thr		
305	310	315
acg gtc agc ctg ctc aag aac ctg ttg gaa ggc gtt cag cac gag gag	1008	
Thr Val Ser Leu Leu Lys Asn Leu Leu Glu Gly Val Gln His Glu Glu		
325	330	335
cag ccg cat gcg ggc tcg tac gag ccg agc tac ctg gaa cgc aat ttg	1056	
Gln Pro His Ala Gly Ser Tyr Glu Pro Ser Tyr Leu Glu Arg Asn Leu		
340	345	350
gtc ggc ctc cac acc tac cac aac atc gcg ttc ctg gag aaa ggc gtc	1104	
Val Gly Leu His Thr Tyr His Asn Ile Ala Phe Leu Glu Lys Gly Val		
355	360	365
aac gcc gaa ggc ggc gtt cct gct tgg gtg cca agg agt ctg gac gac	1152	
Asn Ala Glu Gly Gly Val Pro Ala Trp Val Pro Arg Ser Leu Asp Asp		
370	375	380
ata ttg cac ctg gcc gac gtg aac agc gcg gag gac gag tga	1194	

Ile Leu His Leu Asp Val Asn Ser Ala Glu Asp Glu  
385 390 395

<210> 18

<211> 397

<212> PRT

<213> Saccharopolyspora spinosa

<400> 18

Met Ser Thr Thr His Glu Ile Glu Thr Val Glu Arg Ile Ile Leu Ala  
1 5 10 15

Ala Gly Ser Ser Ala Ala Ser Leu Ala Asp Leu Thr Thr Glu Leu Gly  
20 25 30

Leu Ala Arg Ile Ala Pro Val Leu Ile Asp Glu Ile Leu Phe Arg Ala  
35 40 45

Glu Pro Ala Pro Asp Ile Glu Arg Thr Glu Val Ala Val Gln Ile Thr  
50 55 60

His Arg Gly Glu Thr Val Asp Phe Val Leu Thr Leu Gln Ser Gly Glu  
65 70 75 80

Leu Ile Lys Ala Glu Gln Arg Pro Val Gly Asp Val Pro Leu Arg Ile  
85 90 95

Gly Tyr Glu Leu Thr Asp Leu Ile Ala Glu Leu Phe Gly Pro Gly Ala  
100 105 110

Pro Arg Ala Val Gly Ala Arg Ser Thr Asn Phe Leu Arg Thr Thr Thr  
115 120 125

Ser Gly Ser Ile Pro Gly Pro Ser Glu Leu Ser Asp Gly Phe Gln Ala  
130 135 140

Ile Ser Ala Val Val Ala Gly Cys Gly His Arg Arg Pro Asp Leu Asn  
145 150 155 160

Leu Leu Ala Ser His Tyr Arg Thr Asp Lys Trp Gly Gly Leu His Trp  
165 170 175

Phe Thr Pro Leu Tyr Glu Arg His Leu Gly Glu Phe Arg Asp Arg Pro  
180 185 190

Val Arg Ile Leu Glu Ile Gly Val Gly Gly Tyr Asn Phe Asp Gly Gly  
195 200 205

Gly Gly Glu Ser Leu Lys Met Trp Lys Arg Tyr Phe His Arg Gly Leu  
210 215 220

Val Phe Gly Met Asp Val Phe Asp Lys Ser Phe Leu Asp Gln Gln Arg  
225 230 235 240

Leu Cys Thr Val Arg Ala Asp Gln Ser Lys Pro Glu Glu Leu Ala Ala  
245 250 255

Val Asp Asp Lys Tyr Gly Pro Phe Asp Ile Ile Ile Asp Asp Gly Ser  
260 265 270

His Ile Asn Gly His Val Arg Thr Ser Leu Glu Thr Leu Phe Pro Arg  
275 280 285

Leu Arg Ser Gly Gly Val Tyr Val Ile Glu Asp Leu Trp Thr Thr Tyr  
290 295 300

Ala Pro Gly Phe Gly Gly Gln Ala Gln Cys Pro Ala Ala Pro Gly Thr  
305 310 315 320

Thr Val Ser Leu Leu Lys Asn Leu Leu Glu Gly Val Gln His Glu Glu  
325 330 335

Gln Pro His Ala Gly Ser Tyr Glu Pro Ser Tyr Leu Glu Arg Asn Leu  
340 345 350

Val Gly Leu His Thr Tyr His Asn Ile Ala Phe Leu Glu Lys Gly Val  
355 360 365

Asn Ala Glu Gly Gly Val Pro Ala Trp Val Pro Arg Ser Leu Asp Asp  
370 375 380

Ile Leu His Leu Ala Asp Val Asn Ser Ala Glu Asp Glu  
385 390 395

<210> 19

<211> 900

<212> DNA

<213> *Saccharopolyspora spinosa*

<220>

<221> CDS

<222> (1)..(897)

<223> ORF7; O-Methyltransferase

<400> 19

gtg aac agc aga ggg gcg aac aca cag gca ttt ccg acc gcg gat cag 48  
Val Asn Ser Arg Gly Ala Asn Thr Gln Ala Phe Pro Thr Ala Asp Gln  
1 5 10 15

gtg gag tcc atc ttc gat gcg ttg gcg cac ggg cgt ccc ctg cac cac 96  
Val Glu Ser Ile Phe Asp Ala Leu Ala His Gly Arg Pro Leu His His  
20 25 30

ggt tac tgg gcg ggc ggg tat cgg gag gat gcc ggt gcc aca ccg tgg 144  
Gly Tyr Trp Ala Gly Gly Tyr Arg Glu Asp Ala Gly Ala Thr Pro Trp  
35 40 45

tcg gat gct gcc gac caa ctg acc gac ctg ttc atc gac aag gcc gcg 192  
Ser Asp Ala Ala Asp Gln Leu Thr Asp Leu Phe Ile Asp Lys Ala Ala  
50 55 60

ctc cgt ccc gga gcg cac ctg ttc gac ctg ggc tgc ggc aat ggg cag 240  
Leu Arg Pro Gly Ala His Leu Phe Asp Leu Gly Cys Gly Asn Gly Gln  
65 70 75 80

ccc gta gtc cgt gcg gca tgc gcc agc ggc gtt cga gtc acc gga atc 288  
Pro Val Val Arg Ala Ala Cys Ala Ser Gly Val Arg Val Thr Gly Ile  
85 90 95

acc gtg aac gcc cag cat ctc gcc gcc gcc acc agg ctc gcc aac gag 336  
Thr Val Asn Ala Gln His Leu Ala Ala Thr Arg Leu Ala Asn Glu  
100 105 110

acc gga ctg gcc ggc agt ctt gag ttc gat cta gtc gac ggc gcc cag 384  
Thr Gly Leu Ala Gly Ser Leu Glu Phe Asp Leu Val Asp Gly Ala Gln  
115 120 125

ctg ccc tac ccg gac ggt ttc ttt cag gcc gca tgg gcg atg cag tcc 432  
Leu Pro Tyr Pro Asp Gly Phe Phe Gln Ala Ala Trp Ala Met Gln Ser  
130 135 140

gtc gtg cag atc gtg gac cag gcc gcc gcg atc cgc gag gtc cac cga 480  
Val Val Gln Ile Val Asp Gln Ala Ala Ala Ile Arg Glu Val His Arg  
145 150 155 160

---

atc ctg gaa ccc ggc ggc cgg ttc gtc ctc gga gac atc atc act cgg 528  
Ile Leu Glu Pro Gly Gly Arg Phe Val Leu Gly Asp Ile Ile Thr Arg  
165 170 175

gtt cga ctc ccg gaa gag tac gcg gcg gtt tgg acg ggc acg acc gcc 576  
Val Arg Leu Pro Glu Glu Tyr Ala Ala Val Trp Thr Gly Thr Thr Ala

180

185

190

cat acc ttg aac agc ttc acg gcg ctg gtc agc gaa gcc ggg ttc gag 624  
 His Thr Leu Asn Ser Phe Thr Ala Leu Val Ser Glu Ala Gly Phe Glu  
 195 200 205

att ctc gaa gtc acc gac ctc acg gca cag acc agg tgc atg gtc tcc 672  
 Ile Leu Glu Val Thr Asp Leu Thr Ala Gln Thr Arg Cys Met Val Ser  
 210 215 220

tgg tac gtc gac gag ttg ctc cgg aaa ctc gat gag ctc gcc ggc gtc 720  
 Trp Tyr Val Asp Glu Leu Leu Arg Lys Leu Asp Glu Leu Ala Gly Val  
 225 230 235 240

gag cct gcg gct gtc ggc acc tac cag caa cgc tac ttg gga gac atc 768  
 Glu Pro Ala Ala Val Gly Thr Tyr Gln Gln Arg Tyr Leu Gly Asp Ile  
 245 250 255

gcg gcg aag cac gga ccg gga cca gca cag ctg atc gcc gcg gtt gcg 816  
 Ala Ala Lys His Gly Pro Gly Pro Ala Gln Leu Ile Ala Ala Val Ala  
 260 265 270

gaa tac cgg aaa cat ccg gat tac gcc aga aac gag gaa agc atg ggt 864  
 Glu Tyr Arg Lys His Pro Asp Tyr Ala Arg Asn Glu Glu Ser Met Gly  
 275 280 285

ttc atg ctc ctg cag gct cga aag aag cag tcc tga 900  
 Phe Met Leu Leu Gln Ala Arg Lys Lys Gln Ser  
 290 295

&lt;210&gt; 20

&lt;211&gt; 299

&lt;212&gt; PRT

&lt;213&gt; Saccharopolyspora spinosa

&lt;400&gt; 20

Val Asn Ser Arg Gly Ala Asn Thr Gln Ala Phe Pro Thr Ala Asp Gln  
 1 5 10 15

Val Glu Ser Ile Phe Asp Ala Leu Ala His Gly Arg Pro Leu His His  
 20 25 30

Gly Tyr Trp Ala Gly Gly Tyr Arg Glu Asp Ala Gly Ala Thr Pro Trp  
 35 40 45

Ser Asp Ala Ala Asp Gln Leu Thr Asp Leu Phe Ile Asp Lys Ala Ala  
 50 55 60

Leu Arg Pro Gly Ala His Leu Phe Asp Leu Gly Cys Gly Asn Gly Gln  
65 70 75 80

Pro Val Val Arg Ala Ala Cys Ala Ser Gly Val Arg Val Thr Gly Ile  
85 90 95

Thr Val Asn Ala Gln His Leu Ala Ala Ala Thr Arg Leu Ala Asn Glu  
100 105 110

Thr Gly Leu Ala Gly Ser Leu Glu Phe Asp Leu Val Asp Gly Ala Gln  
115 120 125

Leu Pro Tyr Pro Asp Gly Phe Phe Gln Ala Ala Trp Ala Met Gln Ser  
130 135 140

Val Val Gln Ile Val Asp Gln Ala Ala Ala Ile Arg Glu Val His Arg  
145 150 155 160

Ile Leu Glu Pro Gly Gly Arg Phe Val Leu Gly Asp Ile Ile Thr Arg  
165 170 175

Val Arg Leu Pro Glu Glu Tyr Ala Ala Val Trp Thr Gly Thr Thr Ala  
180 185 190

His Thr Leu Asn Ser Phe Thr Ala Leu Val Ser Glu Ala Gly Phe Glu  
195 200 205

Ile Leu Glu Val Thr Asp Leu Thr Ala Gln Thr Arg Cys Met Val Ser  
210 215 220

Trp Tyr Val Asp Glu Leu Leu Arg Lys Leu Asp Glu Leu Ala Gly Val  
225 230 235 240

Glu Pro Ala Ala Val Gly Thr Tyr Gln Gln Arg Tyr Leu Gly Asp Ile  
245 250 255

Ala Ala Lys His Gly Pro Gly Pro Ala Gln Leu Ile Ala Ala Val Ala  
260 265 270

Glu Tyr Arg Lys His Pro Asp Tyr Ala Arg Asn Glu Glu Ser Met Gly  
275 280 285

Phe Met Leu Leu Gln Ala Arg Lys Lys Gln Ser  
290 295

<210> 21  
 <211> 1167  
 <212> DNA  
 <213> Saccharopolyspora spinosa

<220>  
 <221> CDS  
 <222> (1)..(1164)  
 <223> ORF8; Cyclisierungsenzym

<400> 21

atg gcc tcc gag cac gcc agc ctg gtc ggc gac gat ctg cgg gca ccc 48  
 Met Ala Ser Glu His Ala Ser Leu Val Gly Asp Asp Leu Arg Ala Pro  
 1 5 10 15

gcg gat gat ccc ttc tac cga ccg ccg acg ccg cta ccg ccg ggt gtc 96  
 Ala Asp Asp Pro Phe Tyr Arg Pro Pro Thr Pro Leu Pro Pro Gly Val  
 20 25 30

ccg ggc acg ctc ctc agg gcc cgg ccc gtc tcg gca ctg cgc ggc acg 144  
 Pro Gly Thr Leu Leu Arg Ala Arg Pro Val Ser Ala Leu Arg Gly Thr  
 35 40 45

ggc gaa ccc gtc gca gcc aag gcc tgg caa atc ctc tac cgg tcc aac 192  
 Gly Glu Pro Val Ala Ala Lys Ala Trp Gln Ile Leu Tyr Arg Ser Asn  
 50 55 60

tcc gcc ctt ggc atg ccg aac gcc gtc tcc ggc acc gtt ctg gtg ccg 240  
 Ser Ala Leu Gly Met Pro Asn Ala Val Ser Gly Thr Val Leu Val Pro  
 65 70 75 80

aac atc ccg tgg ccg cgc gaa gat cgc ccc atc atc act ttc gca gtg 288  
 Asn Ile Pro Trp Pro Arg Glu Asp Arg Pro Ile Ile Thr Phe Ala Val  
 85 90 95

ggc acc cac ggc ctc ggt agc caa gtt gcc ccg tcg tac ctg ctt cga 336  
 Gly Thr His Gly Leu Gly Ser Gln Val Ala Pro Ser Tyr Leu Leu Arg  
 100 105 110

acc gga acc gag ccg gag acc gag ctg atc gcc gtg gcc ctc gac cgc 384  
 Thr Gly Thr Glu Pro Glu Thr Glu Leu Ile Ala Val Ala Leu Asp Arg  
 115 120 125

---

ggg tgg gcc gtg gtc atc acc gac tac gaa ggc ctc ggt act cct gga 432  
 Gly Trp Ala Val Val Ile Thr Asp Tyr Glu Gly Leu Gly Thr Pro Gly  
 130 135 140

acc cac acc tac acc gtc ggc agg gcg cag gga cac gcc atg ctc gat 480

Thr His Thr Tyr Thr Val Gly Arg Ala Gln Gly His Ala Met Leu Asp  
 145 150 155 160

gcc gcc cgc gct gcg caa cgg cta ccg ggc tcc ggc ctg acg acc gac 528  
 Ala Ala Arg Ala Ala Gln Arg Leu Pro Gly Ser Gly Leu Thr Thr Asp  
 165 170 175

tgc ccg gtc ggc atc tgg ggc tat gcg cag ggt ggg caa gcg tcg gcc 576  
 Cys Pro Val Gly Ile Trp Gly Tyr Ala Gln Gly Gly Gln Ala Ser Ala  
 180 185 190

ttc gcc ggc gaa ctg cac ccc acc tac gca cct gaa ctg cga atc cgc 624  
 Phe Ala Gly Glu Leu His Pro Thr Tyr Ala Pro Glu Leu Arg Ile Arg  
 195 200 205

gct gcg gcc gca ggt gcg gtg ccg atc gat ctg ctg gac atc atc cac 672  
 Ala Ala Ala Ala Gly Ala Val Pro Ile Asp Leu Leu Asp Ile Ile His  
 210 215 220

cga aat gac ggg gtg ttc acc ggg ccg gtg ctg gcc ggc ctg gtc ggg 720  
 Arg Asn Asp Gly Val Phe Thr Gly Pro Val Leu Ala Gly Leu Val Gly  
 225 230 235 240

cat gcc gct gcc tac ccc gat ctg cca ttc gac gag ctt ctc acc gaa 768  
 His Ala Ala Ala Tyr Pro Asp Leu Pro Phe Asp Glu Leu Leu Thr Glu  
 245 250 255

gcg ggt cgt acc gcc gtt gat caa gtg cgc gag ctc ggt gca ccg gag 816  
 Ala Gly Arg Thr Ala Val Asp Gln Val Arg Glu Leu Gly Ala Pro Glu  
 260 265 270

ctc gtc acc cgc ttc ctc ggc cgc gag ctg agc gac ttc ctc gac act 864  
 Leu Val Thr Arg Phe Leu Gly Arg Glu Leu Ser Asp Phe Leu Asp Thr  
 275 280 285

tcc ggc ctt ttc gag caa cct cga tgg cgc gca cga ctg gcc gaa agc 912  
 Ser Gly Leu Phe Glu Gln Pro Arg Trp Arg Ala Arg Leu Ala Glu Ser  
 290 295 300

gtc gca ggt agg aac ggt ggc ccg gtg gtc ccc acg ctc gtc tac cac 960  
 Val Ala Gly Arg Asn Gly Gly Pro Val Val Pro Thr Leu Val Tyr His  
 305 310 315 320

agc acg gac gac gag atc gtt ccg ttc gca ttc ggc gag cga ctc cgg 1008  
 Ser Thr Asp Asp Glu Ile Val Pro Phe Ala Phe Gly Glu Arg Leu Arg  
 325 330 335

gac agc tac cgc gcg gcg ggt acg cca gtg cgg tgg cat ccg ctc tcc 1056



Asp Ser Tyr Arg Ala Ala Gly Thr Pro Val Arg Trp His Pro Leu Ser  
 340 345 350

gga ttg gct cac ttt ccc gcc gcc ctg gcc agc tcg cga gtg gtc gtc 1104  
 Gly Leu Ala His Phe Pro Ala Ala Leu Ala Ser Ser Arg Val Val Val  
 355 360 365

tcg tgg ttc gac gag cac ttc tcc gag ccg tcc gcg atc agc ggt ccg 1152  
 Ser Trp Phe Asp Glu His Phe Ser Glu Pro Ser Ala Ile Ser Gly Pro  
 370 375 380

cga gat gcc agg tga 1167  
 Arg Asp Ala Arg  
 385

<210> 22

<211> 388

<212> PRT

<213> Saccharopolyspora spinosa

<400> 22

Met Ala Ser Glu His Ala Ser Leu Val Gly Asp Asp Leu Arg Ala Pro  
 1 5 10 15

Ala Asp Asp Pro Phe Tyr Arg Pro Pro Thr Pro Leu Pro Pro Gly Val  
 20 25 30

Pro Gly Thr Leu Leu Arg Ala Arg Pro Val Ser Ala Leu Arg Gly Thr  
 35 40 45

Gly Glu Pro Val Ala Ala Lys Ala Trp Gln Ile Leu Tyr Arg Ser Asn  
 50 55 60

Ser Ala Leu Gly Met Pro Asn Ala Val Ser Gly Thr Val Leu Val Pro  
 65 70 75 80

Asn Ile Pro Trp Pro Arg Glu Asp Arg Pro Ile Ile Thr Phe Ala Val  
 85 90 95

Gly Thr His Gly Leu Gly Ser Gln Val Ala Pro Ser Tyr Leu Leu Arg  
 100 105 110

Thr Gly Thr Glu Pro Glu Thr Glu Leu Ile Ala Val Ala Leu Asp Arg  
 115 120 125

Gly Trp Ala Val Val Ile Thr Asp Tyr Glu Gly Leu Gly Thr Pro Gly  
 130 135 140

Thr His Thr Tyr Thr Val Gly Arg Ala Gln Gly His Ala Met Leu Asp  
145 150 155 160

Ala Ala Arg Ala Ala Gln Arg Leu Pro Gly Ser Gly Leu Thr Thr Asp  
165 170 175

Cys Pro Val Gly Ile Trp Gly Tyr Ala Gln Gly Gly Gln Ala Ser Ala  
180 185 190

Phe Ala Gly Glu Leu His Pro Thr Tyr Ala Pro Glu Leu Arg Ile Arg  
195 200 205

Ala Ala Ala Ala Gly Ala Val Pro Ile Asp Leu Leu Asp Ile Ile His  
210 215 220

Arg Asn Asp Gly Val Phe Thr Gly Pro Val Leu Ala Gly Leu Val Gly  
225 230 235 240

His Ala Ala Ala Tyr Pro Asp Leu Pro Phe Asp Glu Leu Leu Thr Glu  
245 250 255

Ala Gly Arg Thr Ala Val Asp Gln Val Arg Glu Leu Gly Ala Pro Glu  
260 265 270

Leu Val Thr Arg Phe Leu Gly Arg Glu Leu Ser Asp Phe Leu Asp Thr  
275 280 285

Ser Gly Leu Phe Glu Gln Pro Arg Trp Arg Ala Arg Leu Ala Glu Ser  
290 295 300

Val Ala Gly Arg Asn Gly Gly Pro Val Val Pro Thr Leu Val Tyr His  
305 310 315 320

Ser Thr Asp Asp Glu Ile Val Pro Phe Ala Phe Gly Glu Arg Leu Arg  
325 330 335

Asp Ser Tyr Arg Ala Ala Gly Thr Pro Val Arg Trp His Pro Leu Ser  
340 345 350

Gly Leu Ala His Phe Pro Ala Ala Leu Ala Ser Ser Arg Val Val Val  
355 360 365

Ser Trp Phe Asp Glu His Phe Ser Glu Pro Ser Ala Ile Ser Gly Pro  
370 375 380

Arg Asp Ala Arg  
385

<210> 23  
 <211> 1011  
 <212> DNA  
 <213> Saccharopolyspora spinosa

<220>  
 <221> CDS  
 <222> (1)..(1008)  
 <223> ORF9; 2,3-Reduktase

<400> 23  
 atg acc agc tcg atg cga aag ccg gtg cgc atc ggt gtg ctc ggg tgc 48  
 Met Thr Ser Ser Met Arg Lys Pro Val Arg Ile Gly Val Leu Gly Cys  
 1 5 10 15

gct tcc ttc gcg tgg cga cgg atg ctg ccc gcg atg tgc gac gtg gcc 96  
 Ala Ser Phe Ala Trp Arg Arg Met Leu Pro Ala Met Cys Asp Val Ala  
 20 25 30

gaa aca gag gtg gtg gcg gtg gcg agc cgt gat ccg gcg aaa gcc gaa 144  
 Glu Thr Glu Val Val Ala Val Ala Ser Arg Asp Pro Ala Lys Ala Glu  
 35 40 45

cgg ttc gca gcg cga ttc gaa tgc gag gcg gtg ctg ggt tac cag cgg 192  
 Arg Phe Ala Ala Arg Phe Glu Cys Glu Ala Val Leu Gly Tyr Gln Arg  
 50 55 60

ctc ctg gag cgg ccg gac atc gat gcc gtc tac gtg ccg ttg ccg cct 240  
 Leu Leu Glu Arg Pro Asp Ile Asp Ala Val Tyr Val Pro Leu Pro Pro  
 65 70 75 80

ggc atg cat gca gag tgg atc ggc aag gcg ctt gag gca gac aaa cac 288  
 Gly Met His Ala Glu Trp Ile Gly Lys Ala Leu Glu Ala Asp Lys His  
 85 90 95

gtg ctt gcg gag aaa ccg ctg acg acg acg gcg tcc gac acc gct cgc 336  
 Val Leu Ala Glu Lys Pro Leu Thr Thr Thr Ala Ser Asp Thr Ala Arg  
 100 105 110

---

ctg gtc ggg ctg gcc agg agg aag aac ctg ctg ctg cgg gag aat tac 384  
 Leu Val Gly Leu Ala Arg Arg Lys Asn Leu Leu Leu Arg Glu Asn Tyr  
 115 120 125

ctg ttc ctc cac cac ggc cgg cac gac gtg gtc cgc gac ctg ctg caa 432  
 Leu Phe Leu His His Gly Arg His Asp Val Val Arg Asp Leu Leu Gln

130

135

140

tcc ggg gag atc ggt gag ctc cgg gag ttc acc gcc gtg ttc gga att 480  
 Ser Gly Glu Ile Gly Glu Leu Arg Glu Phe Thr Ala Val Phe Gly Ile  
 145 150 155 160

ccg ccg ctt ccc gac acg gac atc cgc tat cgc acc gaa ctc ggt ggc 528  
 Pro Pro Leu Pro Asp Thr Asp Ile Arg Tyr Arg Thr Glu Leu Gly Gly  
 165 170 175

gga gcg ttg ctg gac atc ggt gtc tat ccc gcc cgt gcc gct cgg cac 576  
 Gly Ala Leu Leu Asp Ile Gly Val Tyr Pro Ala Arg Ala Ala Arg His  
 180 185 190

ttt ctc ctc ggt ccg ctc acg gtt ctc ggc gca agc tcg cac gag gcc 624  
 Phe Leu Leu Gly Pro Leu Thr Val Leu Gly Ala Ser Ser His Glu Ala  
 195 200 205

cag gag tcg ggc gtc gac ttg tcg ggc agc gtg ctg ctc caa tcg gaa 672  
 Gln Glu Ser Gly Val Asp Leu Ser Gly Ser Val Leu Leu Gln Ser Glu  
 210 215 220

ggt ggc acc gtt gcc cac ctc gga tac ggt ttc gtg cac cac tac cgc 720  
 Gly Gly Thr Val Ala His Leu Gly Tyr Gly Phe Val His His Tyr Arg  
 225 230 235 240

agc gcg tac gag ctg tgg ggg agt cgt ggg cga atc gtc gtc gac cgg 768  
 Ser Ala Tyr Glu Leu Trp Gly Ser Arg Gly Arg Ile Val Val Asp Arg  
 245 250 255

gcg ttc acg ccg ccc gcc gag tgg cag gcc gtg atc cga atc gag cgg 816  
 Ala Phe Thr Pro Pro Ala Glu Trp Gln Ala Val Ile Arg Ile Glu Arg  
 260 265 270

aag ggc gtt gtc gac gag ttg tcc ttg cca gcg gaa gat cag gtt cgc 864  
 Lys Gly Val Val Asp Glu Leu Ser Leu Pro Ala Glu Asp Gln Val Arg  
 275 280 285

aag gcg gtc acc gcc ttc gca cgc gac atc aga gca ggg aca ggc gtg 912  
 Lys Ala Val Thr Ala Phe Ala Arg Asp Ile Arg Ala Gly Thr Gly Val  
 290 295 300

gac gac cct gcg gtg gcc gga gat tcg ggc gaa tcg atg atc cag cag 960  
 Asp Asp Pro Ala Val Ala Gly Asp Ser Gly Glu Ser Met Ile Gln Gln  
 305 310 315 320

gcc gcg ctg gtg gag gcg atc ggt cag gcc cgt cgg tgc ggg tcc aca 1008  
 Ala Ala Leu Val Glu Ala Ile Gly Gln Ala Arg Arg Cys Gly Ser Thr

325

330

335

tag

1011

&lt;210&gt; 24

&lt;211&gt; 336

&lt;212&gt; PRT

&lt;213&gt; Saccharopolyspora spinosa

&lt;400&gt; 24

Met Thr Ser Ser Met Arg Lys Pro Val Arg Ile Gly Val Leu Gly Cys

1

5

10

15

Ala Ser Phe Ala Trp Arg Arg Met Leu Pro Ala Met Cys Asp Val Ala

20

25

30

Glu Thr Glu Val Val Ala Val Ala Ser Arg Asp Pro Ala Lys Ala Glu

35

40

45

Arg Phe Ala Ala Arg Phe Glu Cys Glu Ala Val Leu Gly Tyr Gln Arg

50

55

60

Leu Leu Glu Arg Pro Asp Ile Asp Ala Val Tyr Val Pro Leu Pro Pro

65

70

75

80

Gly Met His Ala Glu Trp Ile Gly Lys Ala Leu Glu Ala Asp Lys His

85

90

95

Val Leu Ala Glu Lys Pro Leu Thr Thr Thr Ala Ser Asp Thr Ala Arg

100

105

110

Leu Val Gly Leu Ala Arg Arg Lys Asn Leu Leu Leu Arg Glu Asn Tyr

115

120

125

Leu Phe Leu His His Gly Arg His Asp Val Val Arg Asp Leu Leu Gln

130

135

140

Ser Gly Glu Ile Gly Glu Leu Arg Glu Phe Thr Ala Val Phe Gly Ile

145

150

155

160

Pro Pro Leu Pro Asp Thr Asp Ile Arg Tyr Arg Thr Glu Leu Gly Gly

165

170

175

Gly Ala Leu Leu Asp Ile Gly Val Tyr Pro Ala Arg Ala Ala Arg His

180

185

190

Phe Leu Leu Gly Pro Leu Thr Val Leu Gly Ala Ser Ser His Glu Ala

195

200

205

Gln Glu Ser Gly Val Asp Leu Ser Gly Ser Val Leu Leu Gln Ser Glu  
 210 215 220

Gly Gly Thr Val Ala His Leu Gly Tyr Gly Phe Val His His Tyr Arg  
 225 230 235 240

Ser Ala Tyr Glu Leu Trp Gly Ser Arg Gly Arg Ile Val Val Asp Arg  
 245 250 255

Ala Phe Thr Pro Pro Ala Glu Trp Gln Ala Val Ile Arg Ile Glu Arg  
 260 265 270

Lys Gly Val Val Asp Glu Leu Ser Leu Pro Ala Glu Asp Gln Val Arg  
 275 280 285

Lys Ala Val Thr Ala Phe Ala Arg Asp Ile Arg Ala Gly Thr Gly Val  
 290 295 300

Asp Asp Pro Ala Val Ala Gly Asp Ser Gly Glu Ser Met Ile Gln Gln  
 305 310 315 320

Ala Ala Leu Val Glu Ala Ile Gly Gln Ala Arg Arg Cys Gly Ser Thr  
 325 330 335

&lt;210&gt; 25

&lt;211&gt; 1461

&lt;212&gt; DNA

&lt;213&gt; Saccharopolyspora spinosa

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(1458)

&lt;223&gt; ORF10; 2,3-Dehydratase

&lt;400&gt; 25

atg agc agt tct gtc gaa gct gag gca agt gct gct gcg ccg ctc ggc 48  
 Met Ser Ser Ser Val Glu Ala Glu Ala Ser Ala Ala Ala Pro Leu Gly  
 1 5 10 15

agc aac aac acg cgg cgg ttc gtc gac tct gcg ctg agc gct tgc aat 96  
 Ser Asn Asn Thr Arg Arg Phe Val Asp Ser Ala Leu Ser Ala Cys Asn  
 20 25 30

ggc atg att ccg acc acg gag ttc cac tgc tgg ctc gcc gat cgg ctg 144

Gly	Met	Ile	Pro	Thr	Thr	Glu	Phe	His	Cys	Trp	Leu	Ala	Asp	Arg	Leu		
		35					40					45					
ggc	gag	aac	agc	ttc	gag	acc	aat	cgc	atc	ccg	ttc	gac	cgc	ctg	tcg	192	
Gly	Glu	Asn	Ser	Phe	Glu	Thr	Asn	Arg	Ile	Pro	Phe	Asp	Arg	Leu	Ser		
	50					55					60						
aaa	tgg	aaa	ttc	gat	gcc	agc	acg	gag	aac	ctg	gtt	cat	gcc	gac	ggt	240	
Lys	Trp	Lys	Phe	Asp	Ala	Ser	Thr	Glu	Asn	Leu	Val	His	Ala	Asp	Gly		
	65				70				75						80		
agg	ttc	ttc	acg	gta	gaa	ggc	ctg	cag	gtc	gag	acc	aac	tat	ggc	gcg	288	
Arg	Phe	Phe	Thr	Val	Glu	Gly	Leu	Gln	Val	Glu	Thr	Asn	Tyr	Gly	Ala		
			85					90					95				
gca	ccc	agc	tgg	cac	cag	ccg	atc	atc	aac	cag	gct	gaa	gta	ggt	atc	336	
Ala	Pro	Ser	Trp	His	Gln	Pro	Ile	Ile	Asn	Gln	Ala	Glu	Val	Gly	Ile		
			100				105						110				
ctc	ggc	att	ctc	gtc	aag	gag	atc	gac	ggc	gtg	ctg	cac	tgc	ctc	atg	384	
Leu	Gly	Ile	Leu	Val	Lys	Glu	Ile	Asp	Gly	Val	Leu	His	Cys	Leu	Met		
		115					120					125					
tca	gca	aag	atg	gaa	ccg	ggc	aac	gtc	aac	gtc	ctg	cag	ctc	tcg	ccg	432	
Ser	Ala	Lys	Met	Glu	Pro	Gly	Asn	Val	Asn	Val	Leu	Gln	Leu	Ser	Pro		
	130					135					140						
acg	gtt	cag	gca	act	cgg	agc	aac	tac	acg	cag	gca	cac	cgt	ggc	agc	480	
Thr	Val	Gln	Ala	Thr	Arg	Ser	Asn	Tyr	Thr	Gln	Ala	His	Arg	Gly	Ser		
	145				150				155					160			
gtt	ccg	ccc	tat	gtg	gac	tac	ttc	ctc	ggg	cgg	ggc	cgc	ggc	cgc	gtg	528	
Val	Pro	Pro	Tyr	Val	Asp	Tyr	Phe	Leu	Gly	Arg	Gly	Arg	Gly	Arg	Val		
			165					170					175				
ctg	gta	gac	gtg	ctc	cag	tct	gaa	cag	ggg	tcc	tgg	ttc	tac	cgg	aag	576	
Leu	Val	Asp	Val	Leu	Gln	Ser	Glu	Gln	Gly	Ser	Trp	Phe	Tyr	Arg	Lys		
			180					185					190				
cgc	aac	cgg	aac	atg	gtg	gtg	gaa	gtc	cag	gag	gaa	gtg	cca	gtc	ctg	624	
Arg	Asn	Arg	Asn	Met	Val	Val	Glu	Val	Gln	Glu	Glu	Val	Pro	Val	Leu		
	195						200					205					
cca	gac	ttc	tgc	tgg	ttg	acg	ctc	ggc	cag	gtg	ctg	gct	ctc	ctt	cgt	672	
Pro	Asp	Phe	Cys	Trp	Leu	Thr	Leu	Gly	Gln	Val	Leu	Ala	Leu	Leu	Arg		
	210					215					220						
cag	gac	aac	atc	gtc	aac	atg	gac	acc	cgg	acg	gtg	ctg	tct	tgc	atc	720	

Gln Asp Asn Ile Val Asn Met Asp Thr Arg Thr Val Leu Ser Cys Ile  
 225 230 235 240

ccg ttc cac gat tcc gcc acc gga ccc gaa cta gcc gcc tcg gag gag 768  
 Pro Phe His Asp Ser Ala Thr Gly Pro Glu Leu Ala Ala Ser Glu Glu  
 245 250 255

ccc ttc cga cag gcg gtg gcc agg tcg ctc tcg cac ggc atc gat tcg 816  
 Pro Phe Arg Gln Ala Val Ala Arg Ser Leu Ser His Gly Ile Asp Ser  
 260 265 270

tcg agt atc tcc gag gcg gtc ggt tgg ttc gag gaa gcc aag gcc cgc 864  
 Ser Ser Ile Ser Glu Ala Val Gly Trp Phe Glu Glu Ala Lys Ala Arg  
 275 280 285

tac cgc ttg cgg gca acg cgc gtt ccg ctg agc agg gtc gac aag tgg 912  
 Tyr Arg Leu Arg Ala Thr Arg Val Pro Leu Ser Arg Val Asp Lys Trp  
 290 295 300

tat cgc acc gat acc gag atc gcc cac cag gac ggc aag tac ttc gcg 960  
 Tyr Arg Thr Asp Thr Glu Ile Ala His Gln Asp Gly Lys Tyr Phe Ala  
 305 310 315 320

gtg atc gcg gtg tcg gtg tcc gcg acc aat cgt gag gtc gcc agc tgg 1008  
 Val Ile Ala Val Ser Val Ser Ala Thr Asn Arg Glu Val Ala Ser Trp  
 325 330 335

acg cag ccg atg atc gaa ccg cga gaa caa ggt gag atc gca ctg ttg 1056  
 Thr Gln Pro Met Ile Glu Pro Arg Glu Gln Gly Glu Ile Ala Leu Leu  
 340 345 350

gtc aag cgg atc ggc gga gtg ctg cac ggt ttg gtc cac gct cgg gtg 1104  
 Val Lys Arg Ile Gly Gly Val Leu His Gly Leu Val His Ala Arg Val  
 355 360 365

gag gct ggg tat aag tgg act gcg gaa atc gct ccc acg gtc cag tgc 1152  
 Glu Ala Gly Tyr Lys Trp Thr Ala Glu Ile Ala Pro Thr Val Gln Cys  
 370 375 380

agt gtg gcc aac tac caa agc acc ccg tcg aac gac tgg ccg ccg ttc 1200  
 Ser Val Ala Asn Tyr Gln Ser Thr Pro Ser Asn Asp Trp Pro Pro Phe  
 385 390 395 400

ttg gac gac gtg ctc acc gcc gat ccc gaa acc gtg cgg tac gaa tcg 1248  
 Leu Asp Asp Val Leu Thr Ala Asp Pro Glu Thr Val Arg Tyr Glu Ser  
 405 410 415

atc ctg tcc gaa gaa ggc ggt cgg ttc tac cag gcg cag aac agg tac 1296



Ile Leu Ser Glu Glu Gly Gly Arg Phe Tyr Gln Ala Gln Asn Arg Tyr  
 420 425 430

cgg atc atc gag gtg cat gag gac ttc gcg gca cga cct ccc agc gac 1344  
 Arg Ile Ile Glu Val His Glu Asp Phe Ala Ala Arg Pro Pro Ser Asp  
 435 440 445

ttc cgg tgg atg act ttg gga cag ttg ggc gag ctg ctc cgg agc acc 1392  
 Phe Arg Trp Met Thr Leu Gly Gln Leu Gly Glu Leu Leu Arg Ser Thr  
 450 455 460

cac ttc ttg aac atc cag gcg cgc agc ttg gtc gcc tcc ctg cat agc 1440  
 His Phe Leu Asn Ile Gln Ala Arg Ser Leu Val Ala Ser Leu His Ser  
 465 470 475 480

ttg tgg gcg ttg ggg cga tga 1461  
 Leu Trp Ala Leu Gly Arg  
 485

<210> 26  
 <211> 486  
 <212> PRT  
 <213> Saccharopolyspora spinosa

<400> 26  
 Met Ser Ser Ser Val Glu Ala Glu Ala Ser Ala Ala Ala Pro Leu Gly  
 1 5 10 15

Ser Asn Asn Thr Arg Arg Phe Val Asp Ser Ala Leu Ser Ala Cys Asn  
 20 25 30

Gly Met Ile Pro Thr Thr Glu Phe His Cys Trp Leu Ala Asp Arg Leu  
 35 40 45

Gly Glu Asn Ser Phe Glu Thr Asn Arg Ile Pro Phe Asp Arg Leu Ser  
 50 55 60

Lys Trp Lys Phe Asp Ala Ser Thr Glu Asn Leu Val His Ala Asp Gly  
 65 70 75 80

---

Arg Phe Phe Thr Val Glu Gly Leu Gln Val Glu Thr Asn Tyr Gly Ala  
 85 90 95

Ala Pro Ser Trp His Gln Pro Ile Ile Asn Gln Ala Glu Val Gly Ile  
 100 105 110

Leu Gly Ile Leu Val Lys Glu Ile Asp Gly Val Leu His Cys Leu Met

115

120

125

Ser Ala Lys Met Glu Pro Gly Asn Val Asn Val Leu Gln Leu Ser Pro  
130 135 140

Thr Val Gln Ala Thr Arg Ser Asn Tyr Thr Gln Ala His Arg Gly Ser  
145 150 155 160

Val Pro Pro Tyr Val Asp Tyr Phe Leu Gly Arg Gly Arg Gly Arg Val  
165 170 175

Leu Val Asp Val Leu Gln Ser Glu Gln Gly Ser Trp Phe Tyr Arg Lys  
180 185 190

Arg Asn Arg Asn Met Val Val Glu Val Gln Glu Glu Val Pro Val Leu  
195 200 205

Pro Asp Phe Cys Trp Leu Thr Leu Gly Gln Val Leu Ala Leu Leu Arg  
210 215 220

Gln Asp Asn Ile Val Asn Met Asp Thr Arg Thr Val Leu Ser Cys Ile  
225 230 235 240

Pro Phe His Asp Ser Ala Thr Gly Pro Glu Leu Ala Ala Ser Glu Glu  
245 250 255

Pro Phe Arg Gln Ala Val Ala Arg Ser Leu Ser His Gly Ile Asp Ser  
260 265 270

Ser Ser Ile Ser Glu Ala Val Gly Trp Phe Glu Glu Ala Lys Ala Arg  
275 280 285

Tyr Arg Leu Arg Ala Thr Arg Val Pro Leu Ser Arg Val Asp Lys Trp  
290 295 300

Tyr Arg Thr Asp Thr Glu Ile Ala His Gln Asp Gly Lys Tyr Phe Ala  
305 310 315 320

Val Ile Ala Val Ser Val Ser Ala Thr Asn Arg Glu Val Ala Ser Trp  
325 330 335

~~Thr Gln Pro Met Ile Glu Pro Arg Glu Gln Gly Glu Ile Ala Leu Leu~~  
340 345 350

Val Lys Arg Ile Gly Gly Val Leu His Gly Leu Val His Ala Arg Val  
355 360 365

Glu Ala Gly Tyr Lys Trp Thr Ala Glu Ile Ala Pro Thr Val Gln Cys

370

375

380

Ser Val Ala Asn Tyr Gln Ser Thr Pro Ser Asn Asp Trp Pro Pro Phe  
 385 390 395 400

Leu Asp Asp Val Leu Thr Ala Asp Pro Glu Thr Val Arg Tyr Glu Ser  
 405 410 415

Ile Leu Ser Glu Glu Gly Gly Arg Phe Tyr Gln Ala Gln Asn Arg Tyr  
 420 425 430

Arg Ile Ile Glu Val His Glu Asp Phe Ala Ala Arg Pro Pro Ser Asp  
 435 440 445

Phe Arg Trp Met Thr Leu Gly Gln Leu Gly Glu Leu Leu Arg Ser Thr  
 450 455 460

His Phe Leu Asn Ile Gln Ala Arg Ser Leu Val Ala Ser Leu His Ser  
 465 470 475 480

Leu Trp Ala Leu Gly Arg  
 485

&lt;210&gt; 27

&lt;211&gt; 524

&lt;212&gt; DNA

&lt;213&gt; Saccharopolyspora spinosa

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(438)

&lt;223&gt; ORF11; Thioesterase

&lt;400&gt; 27

gtg agc aac gtg tgg ccg gaa aca tgg acg ccg ggg ttt ggc agg tgt 48  
 Val Ser Asn Val Trp Pro Glu Thr Trp Thr Pro Gly Phe Gly Arg Cys  
 1 5 10 15

tca tcg ctg ttg cgt cga ctc gga ttc cgc cgt gac cgg gac gat gcc 96  
 Ser Ser Leu Leu Arg Arg Leu Gly Phe Arg Arg Asp Arg Asp Ala  
 20 25 30

agg cga gtc ccg aag tca gat tct tgt cca gaa tcg tcc aat ggg gtg 144  
 Arg Arg Val Pro Lys Ser Asp Ser Cys Pro Glu Ser Ser Asn Gly Val  
 35 40 45

ttg atc tcc cca gag gtt tgc gct cca acc gat ttc cga cga gga tcg 192  
 Leu Ile Ser Pro Glu Val Cys Ala Pro Thr Asp Phe Arg Arg Gly Ser  
 50 55 60

tgg cgc ccg ctg agc aac gac tac cgt gcg gtc gag aca tac cgc tgt 240  
 Trp Arg Pro Leu Ser Asn Asp Tyr Arg Ala Val Glu Thr Tyr Arg Cys  
 65 70 75 80

gcg cca gga gcg aag gtg ggt tgc ccg atc acc gtg ctg gtg gta gat 288  
 Ala Pro Gly Ala Lys Val Gly Cys Pro Ile Thr Val Leu Val Val Asp  
 85 90 95

gcc gag ccg aag gtc acc ttg gat gag gcg gaa gcc tgg cga gag cac 336  
 Ala Glu Pro Lys Val Thr Leu Asp Glu Ala Glu Ala Trp Arg Glu His  
 100 105 110

acc gag gcc gtg gcc gac gtc cgt gtc ttc tcc ggc ggg cat ttc ttc 384  
 Thr Glu Ala Val Ala Asp Val Arg Val Phe Ser Gly Gly His Phe Phe  
 115 120 125

atg acc gaa cgc cag gac gag gtg ctc gcg gtc ctt acg ggc gga tcg 432  
 Met Thr Glu Arg Gln Asp Glu Val Leu Ala Val Leu Thr Gly Gly Ser  
 130 135 140

ctt cga tgatcctcgc caggccgctg gaccagaccg cgacgccctt gggagccggc 488  
 Leu Arg  
 145

gtgcacatcg tcacggcagt gagggattgg gcatga 524

<210> 28  
 <211> 146  
 <212> PRT  
 <213> Saccharopolyspora spinosa

<400> 28  
 Val Ser Asn Val Trp Pro Glu Thr Trp Thr Pro Gly Phe Gly Arg Cys  
 1 5 10 15

Ser Ser Leu Leu Arg Arg Leu Gly Phe Arg Arg Asp Arg Asp Asp Ala  
 20 25 30

Arg Arg Val Pro Lys Ser Asp Ser Cys Pro Glu Ser Ser Asn Gly Val  
 35 40 45

Leu Ile Ser Pro Glu Val Cys Ala Pro Thr Asp Phe Arg Arg Gly Ser  
 50 55 60

Trp Arg Pro Leu Ser Asn Asp Tyr Arg Ala Val Glu Thr Tyr Arg Cys  
65 70 75 80

Ala Pro Gly Ala Lys Val Gly Cys Pro Ile Thr Val Leu Val Val Asp  
85 90 95

Ala Glu Pro Lys Val Thr Leu Asp Glu Ala Glu Ala Trp Arg Glu His  
100 105 110

Thr Glu Ala Val Ala Asp Val Arg Val Phe Ser Gly Gly His Phe Phe  
115 120 125

Met Thr Glu Arg Gln Asp Glu Val Leu Ala Val Leu Thr Gly Gly Ser  
130 135 140

Leu Arg  
145

<210> 29  
<211> 1320  
<212> DNA  
<213> Saccharopolyspora spinosa

<220>  
<221> CDS  
<222> (1)..(1317)  
<223> ORF12; Glycosyltransferase

<400> 29  
atg cgt gtc ctg ttc acc ccg ctg ccg gcg agt tcg cac ttc ttc aac 48  
Met Arg Val Leu Phe Thr Pro Leu Pro Ala Ser Ser His Phe Phe Asn  
1 5 10 15

ctg gtg ccg ttg gcg tgg gcg ttg cgt gcc gcg ggg cac gag gtc cgt 96  
Leu Val Pro Leu Ala Trp Ala Leu Arg Ala Ala Gly His Glu Val Arg  
20 25 30

gtc gcc atc tgc ccg aat atg gtg tcg atg gtc acc gga gca gga ctc 144  
~~Val Ala Ile Cys Pro Asn Met Val Ser Met Val Thr Gly Ala Gly Leu~~  
35 40 45

acc gcg gtt ccc gtc ggc gac gag ctc gac ctc atc tcc ttg gcg gcc 192  
Thr Ala Val Pro Val Gly Asp Glu Leu Asp Leu Ile Ser Leu Ala Ala  
50 55 60

aag aac gaa gtt ctc ggc agc ggg gtc tcg gac gag aag ggg	240
Lys Asn Glu Leu Val Leu Gly Ser Gly Val Ser Phe Asp Glu Lys Gly	
65 70 75 80	
cgg cat ccg gaa ctc ttc gac gag ctg ctg tca atc aac tcc ggc aga	288
Arg His Pro Glu Leu Phe Asp Glu Leu Leu Ser Ile Asn Ser Gly Arg	
85 90 95	
gac acg gac gcc gtg gag caa ctc cac ctt gtg gat gac cga tcg ctg	336
Asp Thr Asp Ala Val Glu Gln Leu His Leu Val Asp Asp Arg Ser Leu	
100 105 110	
gac gat ctc atg ggg ttc gcc gag aaa tgg cag cct gat ctc gtt gtg	384
Asp Asp Leu Met Gly Phe Ala Glu Lys Trp Gln Pro Asp Leu Val Val	
115 120 125	
tgg gac gct atg gtg tgt tcg ggg cca gtt gtg gcg cga gcg ctc ggc	432
Trp Asp Ala Met Val Cys Ser Gly Pro Val Val Ala Arg Ala Leu Gly	
130 135 140	
gca cga cac gtg cgg atg ctc gtc gcc ctc gat gtg tcg ggg tgg ctg	480
Ala Arg His Val Arg Met Leu Val Ala Leu Asp Val Ser Gly Trp Leu	
145 150 155 160	
cgg tcc ggt ttc ctc gaa tac cag gaa tcg aag ccg cct gag cag cgc	528
Arg Ser Gly Phe Leu Glu Tyr Gln Glu Ser Lys Pro Pro Glu Gln Arg	
165 170 175	
gtc gac ccg ctc ggg acg tgg ctg gga gcg aag ctc gcc aag ttc gga	576
Val Asp Pro Leu Gly Thr Trp Leu Gly Ala Lys Leu Ala Lys Phe Gly	
180 185 190	
gcc acg ttc gat gaa gag atc gtg acg ggc caa gcg acc ata gat ccg	624
Ala Thr Phe Asp Glu Glu Ile Val Thr Gly Gln Ala Thr Ile Asp Pro	
195 200 205	
att cca tcc tgg atg cgc ctg cct gtg gac ttg gac tac atc tcg atg	672
Ile Pro Ser Trp Met Arg Leu Pro Val Asp Leu Asp Tyr Ile Ser Met	
210 215 220	
cgf ttc gtg ccg tac aac ggt ccg gcg gtg ttg ccg gag tgg ttg cgc	720
Arg Phe Val Pro Tyr Asn Gly Pro Ala Val Leu Pro Glu Trp Leu Arg	
225 230 235 240	
gaa cga ccg acg aag ccg cgc gtc tgc atc acg cgc ggg ctg acc aag	768
Glu Arg Pro Thr Lys Pro Arg Val Cys Ile Thr Arg Gly Leu Thr Lys	
245 250 255	

cgg cgg ctg agc agg gtg acc gaa cag tac ggg gag caa agt gac cag	816
Arg Arg Leu Ser Arg Val Thr Glu Gln Tyr Gly Glu Gln Ser Asp Gln	
260 265 270	
 gaa caa gca atg gtg gaa agg ttg ttg cgc ggc gcg gcc agg ctc gac	864
Glu Gln Ala Met Val Glu Arg Leu Leu Arg Gly Ala Ala Arg Leu Asp	
275 280 285	
 gtc gag gtg atc gcc acc ttg tct gac gac gaa gta cgg gag atg ggg	912
Val Glu Val Ile Ala Thr Leu Ser Asp Asp Glu Val Arg Glu Met Gly	
290 295 300	
 gag ttg ccc tcg aac gtc cgg gtc cac gaa tac gta ccg ctc aac gaa	960
Glu Leu Pro Ser Asn Val Arg Val His Glu Tyr Val Pro Leu Asn Glu	
305 310 315 320	
 ctg ctg gag tcg tgt tca gtg atc atc cat cat ggc tcg acg acg acg	1008
Leu Leu Glu Ser Cys Ser Val Ile Ile His His Gly Ser Thr Thr Thr	
325 330 335	
 cag gaa acc gcc acg gtc aac ggc gta ccg cag ttg att ctc cct ggg	1056
Gln Glu Thr Ala Thr Val Asn Gly Val Pro Gln Leu Ile Leu Pro Gly	
340 345 350	
 acc ttc tgg gac gaa tct cgt agg gcg gag ctc cta gcc gat cgg gga	1104
Thr Phe Trp Asp Glu Ser Arg Arg Ala Glu Leu Leu Ala Asp Arg Gly	
355 360 365	
 gcc ggt ctg gtc ctc gac ccc gcg acg ttt acc gaa gac gac gtg cga	1152
Ala Gly Leu Val Leu Asp Pro Ala Thr Phe Thr Glu Asp Asp Val Arg	
370 375 380	
 ggt cag ctg gcc cgc ctg ctc gac gag ccg tcg ttc gct gcc aac gcg	1200
Gly Gln Leu Ala Arg Leu Leu Asp Glu Pro Ser Phe Ala Ala Asn Ala	
385 390 395 400	
 gcg ctg atc cgc cgt gaa atc gag gaa agt ccc agc ccg cac gac atc	1248
Ala Leu Ile Arg Arg Glu Ile Glu Glu Ser Pro Ser Pro His Asp Ile	
405 410 415	
 gtt cca cgt ctg gaa aag cta gtt gcc gaa cgt gag aac cgc cgc act	1296
Val Pro Arg Leu Glu Lys Leu Val Ala Glu Arg Glu Asn Arg Arg Thr	
420 425 430	
 ggg cag tct gat ggc cat ccg tga	1320
Gly Gln Ser Asp Gly His Pro	
435	

<210> 30  
<211> 439  
<212> PRT  
<213> Saccharopolyspora spinosa

<400> 30  
Met Arg Val Leu Phe Thr Pro Leu Pro Ala Ser Ser His Phe Phe Asn  
1 5 10 15  
Leu Val Pro Leu Ala Trp Ala Leu Arg Ala Ala Gly His Glu Val Arg  
20 25 30  
Val Ala Ile Cys Pro Asn Met Val Ser Met Val Thr Gly Ala Gly Leu  
35 40 45  
Thr Ala Val Pro Val Gly Asp Glu Leu Asp Leu Ile Ser Leu Ala Ala  
50 55 60  
Lys Asn Glu Leu Val Leu Gly Ser Gly Val Ser Phe Asp Glu Lys Gly  
65 70 75 80  
Arg His Pro Glu Leu Phe Asp Glu Leu Leu Ser Ile Asn Ser Gly Arg  
85 90 95  
Asp Thr Asp Ala Val Glu Gln Leu His Leu Val Asp Asp Arg Ser Leu  
100 105 110  
Asp Asp Leu Met Gly Phe Ala Glu Lys Trp Gln Pro Asp Leu Val Val  
115 120 125  
Trp Asp Ala Met Val Cys Ser Gly Pro Val Val Ala Arg Ala Leu Gly  
130 135 140  
Ala Arg His Val Arg Met Leu Val Ala Leu Asp Val Ser Gly Trp Leu  
145 150 155 160  
Arg Ser Gly Phe Leu Glu Tyr Gln Glu Ser Lys Pro Pro Glu Gln Arg  
165 170 175  
Val Asp Pro Leu Gly Thr Trp Leu Gly Ala Lys Leu Ala Lys Phe Gly  
180 185 190  
Ala Thr Phe Asp Glu Glu Ile Val Thr Gly Gln Ala Thr Ile Asp Pro  
195 200 205  
Ile Pro Ser Trp Met Arg Leu Pro Val Asp Leu Asp Tyr Ile Ser Met  
210 215 220



Arg Phe Val Pro Tyr Asn Gly Pro Ala Val Leu Pro Glu Trp Leu Arg  
225 230 235 240

Glu Arg Pro Thr Lys Pro Arg Val Cys Ile Thr Arg Gly Leu Thr Lys  
245 250 255

Arg Arg Leu Ser Arg Val Thr Glu Gln Tyr Gly Glu Gln Ser Asp Gln  
260 265 270

Glu Gln Ala Met Val Glu Arg Leu Leu Arg Gly Ala Ala Arg Leu Asp  
275 280 285

Val Glu Val Ile Ala Thr Leu Ser Asp Asp Glu Val Arg Glu Met Gly  
290 295 300

Glu Leu Pro Ser Asn Val Arg Val His Glu Tyr Val Pro Leu Asn Glu  
305 310 315 320

Leu Leu Glu Ser Cys Ser Val Ile Ile His His Gly Ser Thr Thr Thr  
325 330 335

Gln Glu Thr Ala Thr Val Asn Gly Val Pro Gln Leu Ile Leu Pro Gly  
340 345 350

Thr Phe Trp Asp Glu Ser Arg Arg Ala Glu Leu Leu Ala Asp Arg Gly  
355 360 365

Ala Gly Leu Val Leu Asp Pro Ala Thr Phe Thr Glu Asp Asp Val Arg  
370 375 380

Gly Gln Leu Ala Arg Leu Leu Asp Glu Pro Ser Phe Ala Ala Asn Ala  
385 390 395 400

Ala Leu Ile Arg Arg Glu Ile Glu Glu Ser Pro Ser Pro His Asp Ile  
405 410 415

Val Pro Arg Leu Glu Lys Leu Val Ala Glu Arg Glu Asn Arg Arg Thr  
420 425 430

Gly Gln Ser Asp Gly His Pro  
435

<210> 31

<211> 1389

<212> DNA

<213> Saccharopolyspora spinosa

<220>

<221> CDS

<222> (1)..(1386)

<223> ORF13; 3,4-Dehydratase

<400> 31

atg cag agc cgg aaa acc aga gcg ctg ggg aaa ggg cgc gcc aga gtg 48  
Met Gln Ser Arg Lys Thr Arg Ala Leu Gly Lys Gly Arg Ala Arg Val  
1 5 10 15

act tcg tgt gac gac act tgc gct acc gct act gag atg gtg ccg gat 96  
Thr Ser Cys Asp Asp Thr Cys Ala Thr Ala Thr Glu Met Val Pro Asp  
20 25 30

gcc aag gac cgg ata ttg gca tcc gta cgc gat tac cac cgc gaa cag 144  
Ala Lys Asp Arg Ile Leu Ala Ser Val Arg Asp Tyr His Arg Glu Gln  
35 40 45

gaa tcc ccg acc ttc gtg gct gga tcg acg ccg atc cgg cca tcg ggc 192  
Glu Ser Pro Thr Phe Val Ala Gly Ser Thr Pro Ile Arg Pro Ser Gly  
50 55 60

gcc gtg ctc gac gag gac gac cgg gtg gca ctg gtg gaa gcc gcg ctg 240  
Ala Val Leu Asp Glu Asp Asp Arg Val Ala Leu Val Glu Ala Ala Leu  
65 70 75 80

gag ctc cgg atc gcc gcg ggc ggg aat gca cgg cga ttc gag agc gag 288  
Glu Leu Arg Ile Ala Ala Gly Gly Asn Ala Arg Arg Phe Glu Ser Glu  
85 90 95

ttc gcc cgc ttc ttc ggc ctc cgc aag gct cat ctc gtc aac tcc ggt 336  
Phe Ala Arg Phe Phe Gly Leu Arg Lys Ala His Leu Val Asn Ser Gly  
100 105 110

tcg tcg gcc aat ctc ctg gca ctg agt tcg ctt acc tcc ccc aaa ctc 384  
Ser Ser Ala Asn Leu Leu Ala Leu Ser Ser Leu Thr Ser Pro Lys Leu  
115 120 125

ggc gag gca cga ctg cgg ccc ggc gac gaa gtg atc act gcg gcg gtc 432  
Gly Glu Ala Arg Leu Arg Pro Gly Asp Glu Val Ile Thr Ala Ala Val  
130 135 140

ggc ttc ccc acg acg atc aat ccg gcg gtc caa aac gga ctc gtc ccg 480  
Gly Phe Pro Thr Thr Ile Asn Pro Ala Val Gln Asn Gly Leu Val Pro  
145 150 155 160

gta ttc gtc gac gtg gaa ctg ggc acc tac aac gca acg cca gac cgc 528  
 Val Phe Val Asp Val Glu Leu Gly Thr Tyr Asn Ala Thr Pro Asp Arg  
                     165                    170                    175

atc aag gcc gcc gtc acg gaa cgg acg cga gcc atc atg ctg gcg cac 576  
 Ile Lys Ala Ala Val Thr Glu Arg Thr Arg Ala Ile Met Leu Ala His  
                     180                    185                    190

acc ctg ggc aac ccc ttc gcc gct gac gaa atc gcg gag atc gca aaa 624  
 Thr Leu Gly Asn Pro Phe Ala Ala Asp Glu Ile Ala Glu Ile Ala Lys  
                     195                    200                    205

gaa cac gag ctg ttc ctc gtc gaa gac aac tgt gat gcg gtg gga tcc 672  
 Glu His Glu Leu Phe Leu Val Glu Asp Asn Cys Asp Ala Val Gly Ser  
                     210                    215                    220

acc tac cgg gga cgg ctg acc gga acc ttc ggc gac ctg aca acg gtc 720  
 Thr Tyr Arg Gly Arg Leu Thr Gly Thr Phe Gly Asp Leu Thr Thr Val  
                     225                    230                    235                    240

agc ttc tat cct gcc cat cac atc acc agc ggc gag ggt ggc tgc gtg 768  
 Ser Phe Tyr Pro Ala His His Ile Thr Ser Gly Glu Gly Gly Cys Val  
                     245                    250                    255

ttg acc ggc agc ctg gaa ttg gct cgc atc atc gag tcg ctg cgt gac 816  
 Leu Thr Gly Ser Leu Glu Leu Ala Arg Ile Ile Glu Ser Leu Arg Asp  
                     260                    265                    270

tgg gga cgg gat tgc tgg tgc gag ccc ggc gtg gac aac acc tgc cgc 864  
 Trp Gly Arg Asp Cys Trp Cys Glu Pro Gly Val Asp Asn Thr Cys Arg  
                     275                    280                    285

aag agg ttc gac tac cac ctc ggt acc ctt cca ccg ggc tac gac cac 912  
 Lys Arg Phe Asp Tyr His Leu Gly Thr Leu Pro Pro Gly Tyr Asp His  
                     290                    295                    300

aag tac acg ttc tcc cac gtc ggt tac aac ctc aag acc acc gac ctg 960  
 Lys Tyr Thr Phe Ser His Val Gly Tyr Asn Leu Lys Thr Thr Asp Leu  
                     305                    310                    315                    320

cag gcc gca ctt gcg ctg agc cag ttg agc aag att tcc gca ttc ggg 1008  
 Gln Ala Ala Leu Ala Leu Ser Gln Leu Ser Lys Ile Ser Ala Phe Gly  
                     325                    330                    335

tcg gca cgc cgc cgt aac tgg cga cgg ttg cgc gaa ggg ctg tcc ggg 1056  
 Ser Ala Arg Arg Arg Asn Trp Arg Arg Leu Arg Glu Gly Leu Ser Gly  
                     340                    345                    350

ttg ccg ggc ctg ctg ctg ccg gta gcc aca ccg cac agc gac ccg agc 1104  
 Leu Pro Gly Leu Leu Leu Pro Val Ala Thr Pro His Ser Asp Pro Ser  
 355 360 365

tgg ttc ggg ttt gcg atc acc atc agt gcg gac gcc ggg ttc acc cgt 1152  
 Trp Phe Gly Phe Ala Ile Thr Ile Ser Ala Asp Ala Gly Phe Thr Arg  
 370 375 380

gcc gcc ctg gtg aac ttc ctg gaa tcc cgc aac atc ggc acc cga ctg 1200  
 Ala Ala Leu Val Asn Phe Leu Glu Ser Arg Asn Ile Gly Thr Arg Leu  
 385 390 395 400

ctg ttc ggc ggt aac atc acc cgg cac ccg gcc ttc gag cag gtg cgg 1248  
 Leu Phe Gly Gly Asn Ile Thr Arg His Pro Ala Phe Glu Gln Val Arg  
 405 410 415

tac ccg atc gcc gac gcg ctc acc aac agc gac atc gtc acc gac cga 1296  
 Tyr Arg Ile Ala Asp Ala Leu Thr Asn Ser Asp Ile Val Thr Asp Arg  
 420 425 430

acc ttc tgg gtc ggc gtc tac cca ggc ata acg gac caa atg atc gac 1344  
 Thr Phe Trp Val Gly Val Tyr Pro Gly Ile Thr Asp Gln Met Ile Asp  
 435 440 445

tac gtc gtc gaa tca atc gct gaa ttc gtg gcc aag agt tcc tag 1389  
 Tyr Val Val Glu Ser Ile Ala Glu Phe Val Ala Lys Ser Ser  
 450 455 460

<210> 32

<211> 462

<212> PRT

<213> Saccharopolyspora spinosa

<400> 32

Met Gln Ser Arg Lys Thr Arg Ala Leu Gly Lys Gly Arg Ala Arg Val  
 1 5 10 15

Thr Ser Cys Asp Asp Thr Cys Ala Thr Ala Thr Glu Met Val Pro Asp  
 20 25 30

Ala Lys Asp Arg Ile Leu Ala Ser Val Arg Asp Tyr His Arg Glu Gln  
 35 40 45

Glu Ser Pro Thr Phe Val Ala Gly Ser Thr Pro Ile Arg Pro Ser Gly  
 50 55 60

Ala Val Leu Asp Glu Asp Asp Arg Val Ala Leu Val Glu Ala Ala Leu

65		70		75		80									
Glu	Leu	Arg	Ile	Ala	Ala	Gly	Gly	Asn	Ala	Arg	Arg	Phe	Glu	Ser	Glu
				85				90					95		
Phe	Ala	Arg	Phe	Phe	Gly	Leu	Arg	Lys	Ala	His	Leu	Val	Asn	Ser	Gly
			100					105					110		
Ser	Ser	Ala	Asn	Leu	Leu	Ala	Leu	Ser	Ser	Leu	Thr	Ser	Pro	Lys	Leu
			115					120					125		
Gly	Glu	Ala	Arg	Leu	Arg	Pro	Gly	Asp	Glu	Val	Ile	Thr	Ala	Ala	Val
			130					135					140		
Gly	Phe	Pro	Thr	Thr	Ile	Asn	Pro	Ala	Val	Gln	Asn	Gly	Leu	Val	Pro
			145					150				155		160	
Val	Phe	Val	Asp	Val	Glu	Leu	Gly	Thr	Tyr	Asn	Ala	Thr	Pro	Asp	Arg
			165					170						175	
Ile	Lys	Ala	Ala	Val	Thr	Glu	Arg	Thr	Arg	Ala	Ile	Met	Leu	Ala	His
			180					185					190		
Thr	Leu	Gly	Asn	Pro	Phe	Ala	Ala	Asp	Glu	Ile	Ala	Glu	Ile	Ala	Lys
			195					200					205		
Glu	His	Glu	Leu	Phe	Leu	Val	Glu	Asp	Asn	Cys	Asp	Ala	Val	Gly	Ser
			210					215				220			
Thr	Tyr	Arg	Gly	Arg	Leu	Thr	Gly	Thr	Phe	Gly	Asp	Leu	Thr	Thr	Val
			225					230			235			240	
Ser	Phe	Tyr	Pro	Ala	His	His	Ile	Thr	Ser	Gly	Glu	Gly	Gly	Cys	Val
			245					250						255	
Leu	Thr	Gly	Ser	Leu	Glu	Leu	Ala	Arg	Ile	Ile	Glu	Ser	Leu	Arg	Asp
			260					265					270		
Trp	Gly	Arg	Asp	Cys	Trp	Cys	Glu	Pro	Gly	Val	Asp	Asn	Thr	Cys	Arg
			275					280				285			
<del>Lys</del>	<del>Arg</del>	<del>Phe</del>	<del>Asp</del>	<del>Tyr</del>	<del>His</del>	<del>Leu</del>	<del>Gly</del>	<del>Thr</del>	<del>Leu</del>	<del>Pro</del>	<del>Pro</del>	<del>Gly</del>	<del>Tyr</del>	<del>Asp</del>	<del>His</del>
			290					295				300			
Lys	Tyr	Thr	Phe	Ser	His	Val	Gly	Tyr	Asn	Leu	Lys	Thr	Thr	Asp	Leu
			305					310			315			320	
Gln	Ala	Ala	Leu	Ala	Leu	Ser	Gln	Leu	Ser	Lys	Ile	Ser	Ala	Phe	Gly

325

330

335

Ser Ala Arg Arg Arg Asn Trp Arg Arg Leu Arg Glu Gly Leu Ser Gly  
 340 345 350

Leu Pro Gly Leu Leu Leu Pro Val Ala Thr Pro His Ser Asp Pro Ser  
 355 360 365

Trp Phe Gly Phe Ala Ile Thr Ile Ser Ala Asp Ala Gly Phe Thr Arg  
 370 375 380

Ala Ala Leu Val Asn Phe Leu Glu Ser Arg Asn Ile Gly Thr Arg Leu  
 385 390 395 400

Leu Phe Gly Gly Asn Ile Thr Arg His Pro Ala Phe Glu Gln Val Arg  
 405 410 415

Tyr Arg Ile Ala Asp Ala Leu Thr Asn Ser Asp Ile Val Thr Asp Arg  
 420 425 430

Thr Phe Trp Val Gly Val Tyr Pro Gly Ile Thr Asp Gln Met Ile Asp  
 435 440 445

Tyr Val Val Glu Ser Ile Ala Glu Phe Val Ala Lys Ser Ser  
 450 455 460

&lt;210&gt; 33

&lt;211&gt; 1158

&lt;212&gt; DNA

&lt;213&gt; Saccharopolyspora spinosa

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(1155)

&lt;223&gt; ORF14; 4-Aminotransferase

&lt;400&gt; 33

gtg atc aac ctg cac cag ccg atc ctc ggc acc gaa gaa ctc gac gcg 48  
 Val Ile Asn Leu His Gln Pro Ile Leu Gly Thr Glu Glu Leu Asp Ala  
 1 5 10 15

atc gcg gag gtg ttc gcc tcc aac tgg atc ggg ctc ggg ccg cgc acc 96  
 Ile Ala Glu Val Phe Ala Ser Asn Trp Ile Gly Leu Gly Pro Arg Thr  
 20 25 30

cgg acg ttc gag gcc gaa ttc gcc cac cac ctg gga gtg gat ccc gaa 144

Arg Thr Phe Glu Ala Glu Phe Ala His His Leu Gly Val Asp Pro Glu	
35 40 45	
cag gtc gtg ttc ctc aac tcg ggg act gcc gcg ctg ttc ctt acc gtg	192
Gln Val Val Phe Leu Asn Ser Gly Thr Ala Ala Leu Phe Leu Thr Val	
50 55 60	
cag gtg ctc gac ctc ggc cca ggc gac gac gtg gta ctt cct tcg ata	240
Gln Val Leu Asp Leu Gly Pro Gly Asp Asp Val Val Leu Pro Ser Ile	
65 70 75 80	
agc ttc gtg gcg gcg gcc aac gcc atc gca tcc tcc ggt gcc cgc ccg	288
Ser Phe Val Ala Ala Ala Asn Ala Ile Ala Ser Ser Gly Ala Arg Pro	
85 90 95	
gtg ttc tgc gac gtc gac ccc cgg acg ttg aac ccc acg ctg gat gat	336
Val Phe Cys Asp Val Asp Pro Arg Thr Leu Asn Pro Thr Leu Asp Asp	
100 105 110	
gtg gcg agg gcc atc acg ccg gcg acc aag gcc gta ttg ctg ctc cac	384
Val Ala Arg Ala Ile Thr Pro Ala Thr Lys Ala Val Leu Leu Leu His	
115 120 125	
tat gga gga tcg ccg gga gaa gtc acc gcg atc gcc gat ttc tgc cgt	432
Tyr Gly Gly Ser Pro Gly Glu Val Thr Ala Ile Ala Asp Phe Cys Arg	
130 135 140	
gaa aag ggc ctc atg ctc atc gag gac tcc gcc tgc gcg gtg gca tcg	480
Glu Lys Gly Leu Met Leu Ile Glu Asp Ser Ala Cys Ala Val Ala Ser	
145 150 155 160	
tcc gtg cac ggc acc gct tgc gga acc ttt ggt gac ctg gcc acg tgg	528
Ser Val His Gly Thr Ala Cys Gly Thr Phe Gly Asp Leu Ala Thr Trp	
165 170 175	
agt ttc gat gcg atg aag atc ctg gtc acc ggg gat ggg ggc atg ttc	576
Ser Phe Asp Ala Met Lys Ile Leu Val Thr Gly Asp Gly Gly Met Phe	
180 185 190	
tac gcg gcg gat ccg gag ctg gcg cac cgc gca aga cga ctc gcc tac	624
Tyr Ala Ala Asp Pro Glu Leu Ala His Arg Ala Arg Arg Leu Ala Tyr	
195 200 205	
cac ggt ctt gag cag atg agc gga ttc gat tcg gcc aag tct tcc aac	672
His Gly Leu Glu Gln Met Ser Gly Phe Asp Ser Ala Lys Ser Ser Asn	
210 215 220	
cgc tgg tgg gat att cgc gtc gaa gac atc ggc cag cgg ctg atc ggg	720





<400> 34

Val Ile Asn Leu His Gln Pro Ile Leu Gly Thr Glu Glu Leu Asp Ala  
1 5 10 15

Ile Ala Glu Val Phe Ala Ser Asn Trp Ile Gly Leu Gly Pro Arg Thr  
20 25 30

Arg Thr Phe Glu Ala Glu Phe Ala His His Leu Gly Val Asp Pro Glu  
35 40 45

Gln Val Val Phe Leu Asn Ser Gly Thr Ala Ala Leu Phe Leu Thr Val  
50 55 60

Gln Val Leu Asp Leu Gly Pro Gly Asp Asp Val Val Leu Pro Ser Ile  
65 70 75 80

Ser Phe Val Ala Ala Ala Asn Ala Ile Ala Ser Ser Gly Ala Arg Pro  
85 90 95

Val Phe Cys Asp Val Asp Pro Arg Thr Leu Asn Pro Thr Leu Asp Asp  
100 105 110

Val Ala Arg Ala Ile Thr Pro Ala Thr Lys Ala Val Leu Leu Leu His  
115 120 125

Tyr Gly Gly Ser Pro Gly Glu Val Thr Ala Ile Ala Asp Phe Cys Arg  
130 135 140

Glu Lys Gly Leu Met Leu Ile Glu Asp Ser Ala Cys Ala Val Ala Ser  
145 150 155 160

Ser Val His Gly Thr Ala Cys Gly Thr Phe Gly Asp Leu Ala Thr Trp  
165 170 175

Ser Phe Asp Ala Met Lys Ile Leu Val Thr Gly Asp Gly Gly Met Phe  
180 185 190

Tyr Ala Ala Asp Pro Glu Leu Ala His Arg Ala Arg Arg Leu Ala Tyr  
195 200 205

His Gly Leu Glu Gln Met Ser Gly Phe Asp Ser Ala Lys Ser Ser Asn  
210 215 220

Arg Trp Trp Asp Ile Arg Val Glu Asp Ile Gly Gln Arg Leu Ile Gly  
225 230 235 240

Asn Asp Met Thr Ala Ala Leu Gly Ser Val Gln Leu Arg Lys Leu Pro

245

250

255

Glu Phe Ile Asn Arg Arg Arg Glu Ile Ala Thr Gln Tyr Asp Arg Leu  
 260 265 270

Leu Ser Asp Val Pro Gly Val Leu Leu Pro Pro Thr Leu Pro Asp Gly  
 275 280 285

His Val Ser Ser His Tyr Phe Tyr Trp Val Gln Leu Ala Pro Glu Ile  
 290 295 300

Arg Asp Gln Val Ala Gln Gln Met Leu Glu Arg Gly Ile Tyr Thr Ser  
 305 310 315 320

Tyr Arg Tyr Pro Pro Leu His Lys Val Pro Ile Tyr Arg Ala Asp Cys  
 325 330 335

Lys Leu Pro Ser Ala Glu His Ala Cys Arg Arg Thr Leu Leu Leu Pro  
 340 345 350

Leu His Pro Ser Leu Asp Asp Ala Glu Val Arg Thr Val Ala Asp Glu  
 355 360 365

Phe Gln Lys Ala Val Glu His His Ile Ser Gln Arg Ser Pro Leu Arg  
 370 375 380

Lys  
 385

&lt;210&gt; 35

&lt;211&gt; 750

&lt;212&gt; DNA

&lt;213&gt; Saccharopolyspora spinosa

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(747)

&lt;223&gt; ORF15; N-Dimethyltransferase

&lt;400&gt; 35

atg tcg cgc gtg agc gac aca ttc gca gaa acc tcc tcg gta tac agc 48  
 Met Ser Arg Val Ser Asp Thr Phe Ala Glu Thr Ser Ser Val Tyr Ser  
 1 5 10 15

cca gat cat gcc gac atc tac gac gcg atc cac tcc gcg cgt ggc cgg 96  
 Pro Asp His Ala Asp Ile Tyr Asp Ala Ile His Ser Ala Arg Gly Arg

20

25

30

gac tgg gca gcc gag gcc ggg gaa gta gtc cag ctc gta cgc acc agg 144  
 Asp Trp Ala Ala Glu Ala Gly Glu Val Val Gln Leu Val Arg Thr Arg  
 35 40 45

ctg ccc gaa gca cag tcc cta ctc gac gtc gcc tgt ggg acc ggg gcg 192  
 Leu Pro Glu Ala Gln Ser Leu Leu Asp Val Ala Cys Gly Thr Gly Ala  
 50 55 60

cac cta gag cga ttc cgt gcc gaa tac gcg aag gtc gcg ggg ctt gaa 240  
 His Leu Glu Arg Phe Arg Ala Glu Tyr Ala Lys Val Ala Gly Leu Glu  
 65 70 75 80

ctg tcc gat gcg atg cgg gag atc gcg atc aga cga gtc cct gag gta 288  
 Leu Ser Asp Ala Met Arg Glu Ile Ala Ile Arg Arg Val Pro Glu Val  
 85 90 95

ccg att cac atc ggt gac atc cgc gat ttc gac ctc ggc gag cca ttc 336  
 Pro Ile His Ile Gly Asp Ile Arg Asp Phe Asp Leu Gly Glu Pro Phe  
 100 105 110

gac gtc atc acc tgc ctg tgc ttt acc gcg gct tac atg cgg acc gtt 384  
 Asp Val Ile Thr Cys Leu Cys Phe Thr Ala Ala Tyr Met Arg Thr Val  
 115 120 125

gac gac ctg cga cgc gtg acg cgg aac atg gcc cgg cac ctg gcc cct 432  
 Asp Asp Leu Arg Arg Val Thr Arg Asn Met Ala Arg His Leu Ala Pro  
 130 135 140

ggt gga gtc gcg gtc atc gaa ccc tgg tgg ttt ccc gac aag ttc atc 480  
 Gly Gly Val Ala Val Ile Glu Pro Trp Trp Phe Pro Asp Lys Phe Ile  
 145 150 155 160

gac ggg ttc gtc acc gga gcc gtc gcg cac cac ggc gag cgg gtg atc 528  
 Asp Gly Phe Val Thr Gly Ala Val Ala His His Gly Glu Arg Val Ile  
 165 170 175

agc cgg cta tcg cac tcg gtc ctg gag ggc cgt acg agc cgg atg acc 576  
 Ser Arg Leu Ser His Ser Val Leu Glu Gly Arg Thr Ser Arg Met Thr  
 180 185 190

gtc cgc tac aca gtc gcc gaa ccc acc ggg atc cgg gac ttc aca gag 624  
 Val Arg Tyr Thr Val Ala Glu Pro Thr Gly Ile Arg Asp Phe Thr Glu  
 195 200 205

ttc gaa atc ctc tcg ctg ttc act gag gac gag tac acc gcc gcg ctc 672  
 Phe Glu Ile Leu Ser Leu Phe Thr Glu Asp Glu Tyr Thr Ala Ala Leu

210

215

220

gaa gac gca ggg atc cgc gcg gaa tac ctt cct gga gca ccg aac ggc 720  
 Glu Asp Ala Gly Ile Arg Ala Glu Tyr Leu Pro Gly Ala Pro Asn Gly  
 225 230 235 240

cga ggc ctg ttc gtc gga atc cgc aac tga 750  
 Arg Gly Leu Phe Val Gly Ile Arg Asn  
 245

&lt;210&gt; 36

&lt;211&gt; 249

&lt;212&gt; PRT

&lt;213&gt; Saccharopolyspora spinosa

&lt;400&gt; 36

Met Ser Arg Val Ser Asp Thr Phe Ala Glu Thr Ser Ser Val Tyr Ser  
 1 5 10 15

Pro Asp His Ala Asp Ile Tyr Asp Ala Ile His Ser Ala Arg Gly Arg  
 20 25 30

Asp Trp Ala Ala Glu Ala Gly Glu Val Val Gln Leu Val Arg Thr Arg  
 35 40 45

Leu Pro Glu Ala Gln Ser Leu Leu Asp Val Ala Cys Gly Thr Gly Ala  
 50 55 60

His Leu Glu Arg Phe Arg Ala Glu Tyr Ala Lys Val Ala Gly Leu Glu  
 65 70 75 80

Leu Ser Asp Ala Met Arg Glu Ile Ala Ile Arg Arg Val Pro Glu Val  
 85 90 95

Pro Ile His Ile Gly Asp Ile Arg Asp Phe Asp Leu Gly Glu Pro Phe  
 100 105 110

Asp Val Ile Thr Cys Leu Cys Phe Thr Ala Ala Tyr Met Arg Thr Val  
 115 120 125

Asp Asp Leu Arg Arg Val Thr Arg Asn Met Ala Arg His Leu Ala Pro  
 130 135 140

Gly Gly Val Ala Val Ile Glu Pro Trp Trp Phe Pro Asp Lys Phe Ile  
 145 150 155 160

Asp Gly Phe Val Thr Gly Ala Val Ala His His Gly Glu Arg Val Ile

165	170	175
Ser Arg Leu Ser His Ser Val Leu Glu Gly Arg Thr Ser Arg Met Thr		
180	185	190
Val Arg Tyr Thr Val Ala Glu Pro Thr Gly Ile Arg Asp Phe Thr Glu		
195	200	205
Phe Glu Ile Leu Ser Leu Phe Thr Glu Asp Glu Tyr Thr Ala Ala Leu		
210	215	220
Glu Asp Ala Gly Ile Arg Ala Glu Tyr Leu Pro Gly Ala Pro Asn Gly		
225	230	235
240		
Arg Gly Leu Phe Val Gly Ile Arg Asn		
245		

<210> 37  
 <211> 726  
 <212> DNA  
 <213> Saccharopolyspora spinosa

<220>  
 <221> CDS  
 <222> (1)..(723)  
 <223> ORF16; 3,4-Reduktase

<400> 37  
 atg agc gaa cag acg att gca ctg gtc acc ggc gca aac aag gga atc 48  
 Met Ser Glu Gln Thr Ile Ala Leu Val Thr Gly Ala Asn Lys Gly Ile  
 1 5 10 15  
 gga tac gag atc gcg gcc ggg ctc ggc gcg ctg ggg tgg agc gtc gga 96  
 Gly Tyr Glu Ile Ala Ala Gly Leu Gly Ala Leu Gly Trp Ser Val Gly  
 20 25 30  
 atc ggg gca cgg gac cac cag cgc ggg gag gat gcc gtg gcg aaa ttg 144  
 Ile Gly Ala Arg Asp His Gln Arg Gly Glu Asp Ala Val Ala Lys Leu  
 35 40 45

---

cgt gcg gac ggc gtc gat gcg ttc gcg gta tcc ctg gac gtg aca gac 192  
 Arg Ala Asp Gly Val Asp Ala Phe Ala Val Ser Leu Asp Val Thr Asp  
 50 55 60

gac gcg agc gtc gcg gct gct gcg gct ctg ctc gag gag cgc gcc ggc 240  
 Asp Ala Ser Val Ala Ala Ala Ala Ala Leu Leu Glu Glu Arg Ala Gly

65

70

75

80

cgg ctc gat gtg ctg gtt aat aac gcc ggc atc gcc ggg gca tgg ccg 288  
 Arg Leu Asp Val Leu Val Asn Asn Ala Gly Ile Ala Gly Ala Trp Pro  
 85 90 95

gag gag ccc tcg acc gtc aca ccg gcg agc ctc cgg gcg gtg gtg gag 336  
 Glu Glu Pro Ser Thr Val Thr Pro Ala Ser Leu Arg Ala Val Val Glu  
 100 105 110

acc aac gtg atc ggc gtc gtt cgg gtt acc aac gct atg ctg ccg ttg 384  
 Thr Asn Val Ile Gly Val Val Arg Val Thr Asn Ala Met Leu Pro Leu  
 115 120 125

cta cgc cgc tcc gag cgc ccg cgg atc gtc aac cag tcc agc cac gtc 432  
 Leu Arg Arg Ser Glu Arg Pro Arg Ile Val Asn Gln Ser Ser His Val  
 130 135 140

gct tcc ctg acc ttg caa acc acg ccg ggc gtc gac ctc ggc ggg atc 480  
 Ala Ser Leu Thr Leu Gln Thr Thr Pro Gly Val Asp Leu Gly Gly Ile  
 145 150 155 160

agc gga gcc tac tca ccg tcg aag acg ttc ctc aac gcg atc acc atc 528  
 Ser Gly Ala Tyr Ser Pro Ser Lys Thr Phe Leu Asn Ala Ile Thr Ile  
 165 170 175

cag tac gcc aag gaa ctc agc gat acc aac atc aaa atc aac aac gcc 576  
 Gln Tyr Ala Lys Glu Leu Ser Asp Thr Asn Ile Lys Ile Asn Asn Ala  
 180 185 190

tgc ccc ggc tac gtc gcg acc gac ctt aac ggc ttc cac gga acc agc 624  
 Cys Pro Gly Tyr Val Ala Thr Asp Leu Asn Gly Phe His Gly Thr Ser  
 195 200 205

acg ccg gca gac ggt gcc agg atc gcc att cgg ctc gcc acg ctg cca 672  
 Thr Pro Ala Asp Gly Ala Arg Ile Ala Ile Arg Leu Ala Thr Leu Pro  
 210 215 220

gac gac ggc ccg acc gga ggc atg ttc gac gac gcc ggg aat gtg ccc 720  
 Asp Asp Gly Pro Thr Gly Gly Met Phe Asp Asp Ala Gly Asn Val Pro  
 225 230 235 240

tgg tga 726  
 Trp

<210> 38

<211> 241

<212> PRT

<213> Saccharopolyspora spinosa

<400> 38

Met Ser Glu Gln Thr Ile Ala Leu Val Thr Gly Ala Asn Lys Gly Ile  
1 5 10 15

Gly Tyr Glu Ile Ala Ala Gly Leu Gly Ala Leu Gly Trp Ser Val Gly  
20 25 30

Ile Gly Ala Arg Asp His Gln Arg Gly Glu Asp Ala Val Ala Lys Leu  
35 40 45

Arg Ala Asp Gly Val Asp Ala Phe Ala Val Ser Leu Asp Val Thr Asp  
50 55 60

Asp Ala Ser Val Ala Ala Ala Ala Ala Leu Leu Glu Glu Arg Ala Gly  
65 70 75 80

Arg Leu Asp Val Leu Val Asn Asn Ala Gly Ile Ala Gly Ala Trp Pro  
85 90 95

Glu Glu Pro Ser Thr Val Thr Pro Ala Ser Leu Arg Ala Val Val Glu  
100 105 110

Thr Asn Val Ile Gly Val Val Arg Val Thr Asn Ala Met Leu Pro Leu  
115 120 125

Leu Arg Arg Ser Glu Arg Pro Arg Ile Val Asn Gln Ser Ser His Val  
130 135 140

Ala Ser Leu Thr Leu Gln Thr Thr Pro Gly Val Asp Leu Gly Gly Ile  
145 150 155 160

Ser Gly Ala Tyr Ser Pro Ser Lys Thr Phe Leu Asn Ala Ile Thr Ile  
165 170 175

Gln Tyr Ala Lys Glu Leu Ser Asp Thr Asn Ile Lys Ile Asn Asn Ala  
180 185 190

Cys Pro Gly Tyr Val Ala Thr Asp Leu Asn Gly Phe His Gly Thr Ser  
195 200 205

Thr Pro Ala Asp Gly Ala Arg Ile Ala Ile Arg Leu Ala Thr Leu Pro  
210 215 220

Asp Asp Gly Pro Thr Gly Gly Met Phe Asp Asp Ala Gly Asn Val Pro  
225 230 235 240

Trp

<210> 39

<211> 837

<212> DNA

<213> Saccharopolyspora spinosa

<220>

<221> CDS

<222> (1)..(834)

<223> ORF17; Transkriptions-Regulator

<400> 39

atg gag acg cgg gag ttg cgg tac ttc gtt gca gtc gcc gag gag ttg 48

Met Glu Thr Arg Glu Leu Arg Tyr Phe Val Ala Val Ala Glu Glu Leu

1

5

10

15

cac ttc ggc cgg gcc gcc cag cgc ctg ggc atc gcc cag ccg ccg ctg 96

His Phe Gly Arg Ala Ala Gln Arg Leu Gly Ile Ala Gln Pro Pro Leu

20

25

30

tcg cgg acg atc gcc cag ctc gag caa cga ctc gga gtc gtg ttg ctg 144

Ser Arg Thr Ile Ala Gln Leu Glu Gln Arg Leu Gly Val Val Leu Leu

35

40

45

caa cgc acc agc cgc aaa gtc tcg ctc acc gaa gcc ggg gca atg ctg 192

Gln Arg Thr Ser Arg Lys Val Ser Leu Thr Glu Ala Gly Ala Met Leu

50

55

60

ctg acc gaa ggc cgg gcg atc ctc ggc gcg ctg gca gca gcc gag cga 240

Leu Thr Glu Gly Arg Ala Ile Leu Gly Ala Leu Ala Ala Ala Glu Arg

65

70

75

80

cgc acc cag cgt gcc gcg acg agc cag ccc tcg cta gtc ctg gct gcc 288

Arg Thr Gln Arg Ala Ala Thr Ser Gln Pro Ser Leu Val Leu Ala Ala

85

90

95

aag gcc ggc gcc tcc ggt gag ctg ctg gcg aag ttg ctc gac gcg tac 336

Lys Ala Gly Ala Ser Gly Glu Leu Leu Ala Lys Leu Leu Asp Ala Tyr

100

105

110

gcc gcc gag ccg gga gcc gtg gcc gtc gac ctg ctg ctc tgc gaa tcc 384

Ala Ala Glu Pro Gly Ala Val Ala Val Asp Leu Leu Cys Glu Ser

115

120

125



cag ccc cag aaa acg ctg cat gac ggc cgg gcc gac gtg gcg ctg ttg	432
Gln Pro Gln Lys Thr Leu His Asp Gly Arg Ala Asp Val Ala Leu Leu	
130 135 140	
cat caa ccc ttc gac ccg acg gcc gaa ctc gac atc gaa att ctg aac	480
His Gln Pro Phe Asp Pro Thr Ala Glu Leu Asp Ile Glu Ile Leu Asn	
145 150 155 160	
acc gag caa caa gtc gcc att ctt ccg acc tcg cat ccg ctt gcc agc	528
Thr Glu Gln Gln Val Ala Ile Leu Pro Thr Ser His Pro Leu Ala Ser	
165 170 175	
gag ccc cat gta cgg atg gcg gat gtc agc tca ctg ccg gat ctc ccg	576
Glu Pro His Val Arg Met Ala Asp Val Ser Ser Leu Pro Asp Leu Pro	
180 185 190	
ctt gcg cgc tgg ccc ggc ccc gac ggc gtc tat cca gat ggc ccc ggc	624
Leu Ala Arg Trp Pro Gly Pro Asp Gly Val Tyr Pro Asp Gly Pro Gly	
195 200 205	
gtg gaa gta cgc aac cag acg caa ctg ttc caa atg atc gca ctc ggc	672
Val Glu Val Arg Asn Gln Thr Gln Leu Phe Gln Met Ile Ala Leu Gly	
210 215 220	
cgc act acc gtg gtc atg ccc gaa tcc agt cgc gtc aac ctg ctc gaa	720
Arg Thr Thr Val Val Met Pro Glu Ser Ser Arg Val Asn Leu Leu Glu	
225 230 235 240	
ggc ctc gcc gcc gta ccg gtt cta gac gcg ccg gac gtg acg aca gtc	768
Gly Leu Ala Ala Val Pro Val Leu Asp Ala Pro Asp Val Thr Thr Val	
245 250 255	
atc gcc tgg ccg ccc cac agc cgc tcc cga gca ctc gcc ggc ttg gtc	816
Ile Ala Trp Pro Pro His Ser Arg Ser Arg Ala Leu Ala Gly Leu Val	
260 265 270	
cgc gtg gcc aca ctc ctc taa	837
Arg Val Ala Thr Leu Leu	
275	

---

<210> 40  
 <211> 278  
 <212> PRT  
 <213> Saccharopolyspora spinosa

<400> 40

Met Glu Thr Arg Gln Leu Arg Tyr Phe Val Ala Val Ala Leu Glu Leu  
1 5 10 15

His Phe Gly Arg Ala Ala Gln Arg Leu Gly Ile Ala Gln Pro Pro Leu  
20 25 30

Ser Arg Thr Ile Ala Gln Leu Glu Gln Arg Leu Gly Val Val Leu Leu  
35 40 45

Gln Arg Thr Ser Arg Lys Val Ser Leu Thr Glu Ala Gly Ala Met Leu  
50 55 60

Leu Thr Glu Gly Arg Ala Ile Leu Gly Ala Leu Ala Ala Ala Glu Arg  
65 70 75 80

Arg Thr Gln Arg Ala Ala Thr Ser Gln Pro Ser Leu Val Leu Ala Ala  
85 90 95

Lys Ala Gly Ala Ser Gly Glu Leu Leu Ala Lys Leu Leu Asp Ala Tyr  
100 105 110

Ala Ala Glu Pro Gly Ala Val Ala Val Asp Leu Leu Leu Cys Glu Ser  
115 120 125

Gln Pro Gln Lys Thr Leu His Asp Gly Arg Ala Asp Val Ala Leu Leu  
130 135 140

His Gln Pro Phe Asp Pro Thr Ala Glu Leu Asp Ile Glu Ile Leu Asn  
145 150 155 160

Thr Glu Gln Gln Val Ala Ile Leu Pro Thr Ser His Pro Leu Ala Ser  
165 170 175

Glu Pro His Val Arg Met Ala Asp Val Ser Ser Leu Pro Asp Leu Pro  
180 185 190

Leu Ala Arg Trp Pro Gly Pro Asp Gly Val Tyr Pro Asp Gly Pro Gly  
195 200 205

Val Glu Val Arg Asn Gln Thr Gln Leu Phe Gln Met Ile Ala Leu Gly  
210 215 220

Arg Thr Thr Val Val Met Pro Glu Ser Ser Arg Val Asn Leu Leu Glu  
225 230 235 240

Gly Leu Ala Ala Val Pro Val Leu Asp Ala Pro Asp Val Thr Thr Val  
245 250 255

Ile Ala Trp Pro Pro His Ser Arg Ser Arg Ala Leu Ala Gly Leu Val  
 260 265 270

Arg Val Ala Thr Leu Leu  
 275

<210> 41  
 <211> 7788  
 <212> DNA  
 <213> Saccharopolyspora spinosa

<220>  
 <221> CDS  
 <222> (1)..(7785)  
 <223> ORF18; Polyketidsynthase

<400> 41  
 atg agc gaa gcc ggg aac ctg ata gcc gtc atc gga ctg tcc tgc cgc 48  
 Met Ser Glu Ala Gly Asn Leu Ile Ala Val Ile Gly Leu Ser Cys Arg  
 1 5 10 15

cta ccc cag gcg cct gac ccg gct tcc ttc tgg cgg ttg ctg cgc acc 96  
 Leu Pro Gln Ala Pro Asp Pro Ala Ser Phe Trp Arg Leu Leu Arg Thr  
 20 25 30

gga acg gac gcc atc acc acg gtc ccg gaa ggg cgg tgg ggc gac ccg 144  
 Gly Thr Asp Ala Ile Thr Thr Val Pro Glu Gly Arg Trp Gly Asp Pro  
 35 40 45

ttg cct ggt cgg gat gcg ccc aag ggc ccg gaa tgg ggt ggt ttc ctg 192  
 Leu Pro Gly Arg Asp Ala Pro Lys Gly Pro Glu Trp Gly Gly Phe Leu  
 50 55 60

gct gat gtc gac tgc ttc gat ccc gag ttc ttc ggg atc tcg ccg cga 240  
 Ala Asp Val Asp Cys Phe Asp Pro Glu Phe Phe Gly Ile Ser Pro Arg  
 65 70 75 80

gaa gcg gca gcc gtg gat ccc cag cag agg ctg gct ctg gag ctc gcc 288  
 Glu Ala Ala Ala Val Asp Pro Gln Gln Arg Leu Ala Leu Glu Leu Ala  
 85 90 95

tgg gag gca ctc gaa gac gcc ggt atc ccc gcc ggc gag ctg cgc ggt 336  
 Trp Glu Ala Leu Glu Asp Ala Gly Ile Pro Ala Gly Glu Leu Arg Gly  
 100 105 110

act gcc gcc ggt gtg ttc atg ggg gcg atc tct gac gac tac gcc gcc 384

Thr Ala Ala Gly Val Phe Met Gly Ala Ile Ser Asp Asp Tyr Ala Ala  
 115 120 125

ctg ctg cgc gag agc ccg ccg gaa gtg gct gcg cag tac cgc ctc acc 432  
 Leu Leu Arg Glu Ser Pro Pro Glu Val Ala Ala Gln Tyr Arg Leu Thr  
 130 135 140

ggc acc cat cga agc ctg atc gcc aac cgc gtg tcc tat gtg ctc ggc 480  
 Gly Thr His Arg Ser Leu Ile Ala Asn Arg Val Ser Tyr Val Leu Gly  
 145 150 155 160

ctg cgc ggg cca agc ctg acg gtg gat tca ggt cag tcc tcg tcc ctg 528  
 Leu Arg Gly Pro Ser Leu Thr Val Asp Ser Gly Gln Ser Ser Ser Leu  
 165 170 175

gtc ggc gtg cat ctc gcc agc gag agc ctg cga cgg ggt gag tgc acg 576  
 Val Gly Val His Leu Ala Ser Glu Ser Leu Arg Arg Gly Glu Cys Thr  
 180 185 190

atc gca ctc gcc ggc ggc gtg aac ctc aac ctg gcc gcc gag agc aac 624  
 Ile Ala Leu Ala Gly Gly Val Asn Leu Asn Leu Ala Ala Glu Ser Asn  
 195 200 205

agc gct ctg atg gac ttc ggc gcg ctc tcc ccg gac ggt cgc tgc ttc 672  
 Ser Ala Leu Met Asp Phe Gly Ala Leu Ser Pro Asp Gly Arg Cys Phe  
 210 215 220

acc ttc gat gtg cgg gcg aac ggt tac gtc cgt ggt gag ggc ggc ggc 720  
 Thr Phe Asp Val Arg Ala Asn Gly Tyr Val Arg Gly Glu Gly Gly Gly  
 225 230 235 240

ctt gtc gtg ctg aag aag gcc gat cag gcg cac gcc gat ggc gac cgg 768  
 Leu Val Val Leu Lys Lys Ala Asp Gln Ala His Ala Asp Gly Asp Arg  
 245 250 255

atc tac tgc ctc atc cgc ggc agc gcg gtc aac aac gat ggg ggc ggt 816  
 Ile Tyr Cys Leu Ile Arg Gly Ser Ala Val Asn Asn Asp Gly Gly Gly  
 260 265 270

gcc ggg ctc acc gtt ccg gcg gcg gac gcc cag gcg gag ctg ctg cgc 864  
 Ala Gly Leu Thr Val Pro Ala Ala Asp Ala Gln Ala Glu Leu Leu Arg  
 275 280 285

cag gca tac cgg aac gcg ggc gtc gac ccg gcc gcc gtg cag tat gtc 912  
 Gln Ala Tyr Arg Asn Ala Gly Val Asp Pro Ala Ala Val Gln Tyr Val  
 290 295 300

gag ctc cac ggc agc gcg acc agg gtc ggg gat ccc gtc gaa gca gca 960



Pro Lys Leu Leu Ala Gly Leu Lys Ala Leu Ala Asp Gly Asp Asp Ala	
500	510
ccc acg ctg tgc cac ggc gcg act tcc ggc gag cgg gca gcg gtc ttc	1584
Pro Thr Leu Cys His Gly Ala Thr Ser Gly Glu Arg Ala Ala Val Phe	
515	520
gtc ttt ccc gga cag ggc agc cag tgg atc ggg atg ggt agg cag ctg	1632
Val Phe Pro Gly Gln Gly Ser Gln Trp Ile Gly Met Gly Arg Gln Leu	
530	540
ctc gaa acc tcc gag gtt ttc gcg gcg tcg atg tcg gac tgc gcc gac	1680
Leu Glu Thr Ser Glu Val Phe Ala Ala Ser Met Ser Asp Cys Ala Asp	
545	555
gca ttg gcg ccg cac ctg gat tgg tcc ctg ctg gat gtg ctg cgc aac	1728
Ala Leu Ala Pro His Leu Asp Trp Ser Leu Leu Asp Val Leu Arg Asn	
565	570
gcg gcc ggc gct gcg cac ctt gac cac gac gat gtc gtc cag ccc gcg	1776
Ala Ala Gly Ala Ala His Leu Asp His Asp Asp Val Val Gln Pro Ala	
580	590
ctg ttc gcc atc atg gtc tcg ctc gcg gag ctc tgg cgt tcg tgg ggc	1824
Leu Phe Ala Ile Met Val Ser Leu Ala Glu Leu Trp Arg Ser Trp Gly	
595	605
gtg cgt ccg gtg gcg gtc gtc ggg cac tcg cag ggg gag atc gcg gcg	1872
Val Arg Pro Val Ala Val Val Gly His Ser Gln Gly Glu Ile Ala Ala	
610	620
gcc tgc gtc gcc ggg gcc ctg tcc gtc cgc gat gcc gcc agg gtg gtg	1920
Ala Cys Val Ala Gly Ala Leu Ser Val Arg Asp Ala Ala Arg Val Val	
625	635
gcg gtg cgc agc agg ctt ctg acg gcg ctg gcc ggc agt ggc gcg atg	1968
Ala Val Arg Ser Arg Leu Leu Thr Ala Leu Ala Gly Ser Gly Ala Met	
645	655
gcc tcg ttg cag cat ccc gcc gaa gag gtg cgg caa atc ctg ttg ccc	2016
Ala Ser Leu Gln His Pro Ala Glu Glu Val Arg Gln Ile Leu Leu Pro	
660	670
tgg cgc gat cgg atc ggc gtg gcg ggg gtg aac gga ccg tcg tcg acc	2064
Trp Arg Asp Arg Ile Gly Val Ala Gly Val Asn Gly Pro Ser Ser Thr	
675	685
ctg gtg tca ggg gac cgg gag gcg atg gcg gaa ctg ctg gcc gag tgc	2112

Leu Val Ser Gly Asp Arg Glu Ala Met Ala Glu Leu Leu Ala Glu Cys	
690 695 700	
gca gac cga gag ctc cgg atg cgc cgg att ccc gtt gaa tac gcc tcc	2160
Ala Asp Arg Glu Leu Arg Met Arg Arg Ile Pro Val Glu Tyr Ala Ser	
705 710 715 720	
cat tcg cct cac atc gag gtt gtc cgg gat gag ctg ctg ggg ctg ttg	2208
His Ser Pro His Ile Glu Val Val Arg Asp Glu Leu Leu Gly Leu Leu	
725 730 735	
gcg ccg gtc gaa ccc agg acg gga agc atc ccg atc tat tcg acg acg	2256
Ala Pro Val Glu Pro Arg Thr Gly Ser Ile Pro Ile Tyr Ser Thr Thr	
740 745 750	
acc ggg gac ctg ctg gac cgg ccg atg gac gcc gac tac tgg tac cgc	2304
Thr Gly Asp Leu Leu Asp Arg Pro Met Asp Ala Asp Tyr Trp Tyr Arg	
755 760 765	
aac ctt cgt caa ccg gtg ctg ttc gaa gcg gcc gtc gag gcc ctg ttg	2352
Asn Leu Arg Gln Pro Val Leu Phe Glu Ala Ala Val Glu Ala Leu Leu	
770 775 780	
aag cgg ggg tac gac gca ttc atc gag atc agc cca cac ccg gtg ctg	2400
Lys Arg Gly Tyr Asp Ala Phe Ile Glu Ile Ser Pro His Pro Val Leu	
785 790 795 800	
act gcg aac atc cag gaa acc gcc gtg cga gca ggg cgg gag gta gtg	2448
Thr Ala Asn Ile Gln Glu Thr Ala Val Arg Ala Gly Arg Glu Val Val	
805 810 815	
gcg ctc ggg aca ctc cgc cgc ggc gaa ggt ggc atg cgg cag gcg ctg	2496
Ala Leu Gly Thr Leu Arg Arg Gly Glu Gly Gly Met Arg Gln Ala Leu	
820 825 830	
acg tcg ctg gcc aga gca cac gta cac gga gtg gcc gcg gac tgg cac	2544
Thr Ser Leu Ala Arg Ala His Val His Gly Val Ala Ala Asp Trp His	
835 840 845	
gcg gtc ttc gcc ggt acc ggg gcg cag cgg gtc gac ctg ccg acg tac	2592
Ala Val Phe Ala Gly Thr Gly Ala Gln Arg Val Asp Leu Pro Thr Tyr	
850 855 860	
gcc ttt cag cga cag cgc tac tgg ctg gac gcg aag ctt ccc gac gtc	2640
Ala Phe Gln Arg Gln Arg Tyr Trp Leu Asp Ala Lys Leu Pro Asp Val	
865 870 875 880	
gcc atg ccc gag agc gac gtg tcg acg gcg ttg cgg gaa aag ctg cgg	2688

Ala Met Pro Glu Ser Asp Val Ser Thr Ala Leu Arg Glu Lys Leu Arg  
 885 890 895

tct tcg ccg agg gcg gac gtg gac tcg acg acc ctc acg atg atc cgg 2736  
 Ser Ser Pro Arg Ala Asp Val Asp Ser Thr Thr Leu Thr Met Ile Arg  
 900 905 910

gca cag gca gcc gtg gtc ctc ggc cac tcc gat ccg aaa gag gtg gac 2784  
 Ala Gln Ala Ala Val Val Leu Gly His Ser Asp Pro Lys Glu Val Asp  
 915 920 925

ccg gat ccg acg ttc aag gac ctg ggc ttc gat tcc tcg atg gtg gtc 2832  
 Pro Asp Arg Thr Phe Lys Asp Leu Gly Phe Asp Ser Ser Met Val Val  
 930 935 940

gag ctg tgc gac cgc cta aac gcc gcc aca ggt ctg cga ctc gca ccg 2880  
 Glu Leu Cys Asp Arg Leu Asn Ala Ala Thr Gly Leu Arg Leu Ala Pro  
 945 950 955 960

agc gtc gtt ttc gac tgt cct acg ccg gac aag ctc gcc cgc cag gta 2928  
 Ser Val Val Phe Asp Cys Pro Thr Pro Asp Lys Leu Ala Arg Gln Val  
 965 970 975

ccg acg ttg ttg ttg ggc gag ccg gct ccc atg acg tca cac ccg ccg 2976  
 Arg Thr Leu Leu Leu Gly Glu Pro Ala Pro Met Thr Ser His Arg Pro  
 980 985 990

gac tcc gat gcg gac gag cct atc gcc gtg atc ggg atg ggc tgt ccg 3024  
 Asp Ser Asp Ala Asp Glu Pro Ile Ala Val Ile Gly Met Gly Cys Arg  
 995 1000 1005

ttt ccg ggt ggg gtg tcc tcg ccc gag gag ttg tgg cag ttg gtc gcc 3072  
 Phe Pro Gly Gly Val Ser Ser Pro Glu Glu Leu Trp Gln Leu Val Ala  
 1010 1015 1020

gct ggg ccg gac gtc gtg tcc gag ttc ccg gct gac cga ggt tgg gac 3120  
 Ala Gly Arg Asp Val Val Ser Glu Phe Pro Ala Asp Arg Gly Trp Asp  
 1025 1030 1035 1040

ctg gag cgt gcg ggg aca tcg cac gtg cgc gcc ggc ggg ttc ttg cat 3168  
 Leu Glu Arg Ala Gly Thr Ser His Val Arg Ala Gly Gly Phe Leu His  
 1045 1050 1055

ggc gcc ccg gat ttt gac ccc ggg ttc ttc ccg att tcg ccg cgc gag 3216  
 Gly Ala Pro Asp Phe Asp Pro Gly Phe Phe Arg Ile Ser Pro Arg Glu  
 1060 1065 1070

gcg ttg gcg atg gat cca cag cag ccg ttg ctg ctg gaa atc gcc tgg 3264



Ala Leu Ala Met Asp Pro Gln Gln Arg Leu Leu Leu Glu Ile Ala Trp  
 1075 1080 1085

gaa gca gtc gaa cga ggc ggg atc aac ccg cag cac ctg cac gga agt 3312  
 Glu Ala Val Glu Arg Gly Gly Ile Asn Pro Gln His Leu His Gly Ser  
 1090 1095 1100

caa acc ggg gtc ttc gtc ggc gcg acc tcc ctg gac tac ggg cca cgc 3360  
 Gln Thr Gly Val Phe Val Gly Ala Thr Ser Leu Asp Tyr Gly Pro Arg  
 1105 1110 1115 1120

ctg cac gaa gcg tcc gag gag gcg gcc ggg tac gtg ctc acc ggc agc 3408  
 Leu His Glu Ala Ser Glu Glu Ala Ala Gly Tyr Val Leu Thr Gly Ser  
 1125 1130 1135

acc acg agt gtg gcg tcg ggt cgg gtt gcg tat tcg ttc ggg ttc gag 3456  
 Thr Thr Ser Val Ala Ser Gly Arg Val Ala Tyr Ser Phe Gly Phe Glu  
 1140 1145 1150

ggc cct gcg gtg acg gtg gat acg gcg tgt tcg tcg tcg ttg gtg gcc 3504  
 Gly Pro Ala Val Thr Val Asp Thr Ala Cys Ser Ser Ser Leu Val Ala  
 1155 1160 1165

ctg cat ttg gcg tgt cag tcg ttg cgt tcg ggt gag tgt gat ctg gcg 3552  
 Leu His Leu Ala Cys Gln Ser Leu Arg Ser Gly Glu Cys Asp Leu Ala  
 1170 1175 1180

ttg gcc ggt ggt gtg acc gtg atg gcc acg ccg ggg atg ttc gtg gag 3600  
 Leu Ala Gly Gly Val Thr Val Met Ala Thr Pro Gly Met Phe Val Glu  
 1185 1190 1195 1200

ttt tcg cgg cag cgt ggt ttg gcg ccg gat ggg cgg tgc aag tcg ttc 3648  
 Phe Ser Arg Gln Arg Gly Leu Ala Pro Asp Gly Arg Cys Lys Ser Phe  
 1205 1210 1215

gcg gag gcc gcc gac ggc acc ggc tgg tcc gag ggt gct ggc ctg gtt 3696  
 Ala Glu Ala Ala Asp Gly Thr Gly Trp Ser Glu Gly Ala Gly Leu Val  
 1220 1225 1230

cta ctg gag cgg ttg tcg gat gcc cgg cgg aat ggg cat gag gtg ctg 3744  
 Leu Leu Glu Arg Leu Ser Asp Ala Arg Arg Asn Gly His Glu Val Leu  
 1235 1240 1245

gcg gtt gtt cgt ggt agt gcg gtg aat cag gac ggt gcg tcg aat ggt 3792  
 Ala Val Val Arg Gly Ser Ala Val Asn Gln Asp Gly Ala Ser Asn Gly  
 1250 1255 1260

ttg acc gcg ccg aat ggt tcg tcg cag cag ccg gtg att gcc cag gca 3840

Leu Thr Ala Pro A	ly Ser Ser Gln Gln Arg Val	Ile Ala Gln Ala	
1265	1270	1275	1280
ttg gcg agt gcg ggg ttg tcg gtg tcc gat gtg gat gct gtg gag gcg			3888
Leu Ala Ser Ala Gly Leu Ser Val Ser Asp Val Asp Ala Val Glu Ala			
1285	1290	1295	
cat ggg acg ggc acg cgg ctt ggt gat ccg atc gag gcg cag gcg ctg			3936
His Gly Thr Gly Thr Arg Leu Gly Asp Pro Ile Glu Ala Gln Ala Leu			
1300	1305	1310	
atc gcc acc tac ggc cag ggc cgg ctt ccg gaa cgg cca ttg tgg ttg			3984
Ile Ala Thr Tyr Gly Gln Gly Arg Leu Pro Glu Arg Pro Leu Trp Leu			
1315	1320	1325	
ggc tcg atg aag tcg aac atc ggt cac gcg cag gca gct gcg ggg ata			4032
Gly Ser Met Lys Ser Asn Ile Gly His Ala Gln Ala Ala Ala Gly Ile			
1330	1335	1340	
gcc ggc gtc atg aag atg gtg atg gcg atg cgg cac ggg cag cta ccg			4080
Ala Gly Val Met Lys Met Val Met Ala Met Arg His Gly Gln Leu Pro			
1345	1350	1355	1360
cgc acg ttg cac gtg gat gag ccg act tct ggg gtg gat tgg tcg gcg			4128
Arg Thr Leu His Val Asp Glu Pro Thr Ser Gly Val Asp Trp Ser Ala			
1365	1370	1375	
ggg acg gtt caa ctc ctt acg gag aac acg ccc tgg ccc ggg agt ggt			4176
Gly Thr Val Gln Leu Leu Thr Glu Asn Thr Pro Trp Pro Gly Ser Gly			
1380	1385	1390	
cgt gtt cgt cgg gtg ggg gtg tcg tcg ttc ggg atc agt ggt act aac			4224
Arg Val Arg Arg Val Gly Val Ser Ser Phe Gly Ile Ser Gly Thr Asn			
1395	1400	1405	
gcg cac gtc atc ctc gaa cag ccc ccg gga gtg ccg agt cag tct gcg			4272
Ala His Val Ile Leu Glu Gln Pro Pro Gly Val Pro Ser Gln Ser Ala			
1410	1415	1420	
ggg ccg ggt tcg ggc tct gtc gtg gat gtt ccg gtg gtg ccg tgg atg			4320
Gly Pro Gly Ser Gly Ser Val Val Asp Val Pro Val Val Pro Trp Met			
1425	1430	1435	1440
gtg tcg ggc aaa aca ccc gaa gcg cta tcc gcg cag gca acg gcg ttg			4368
Val Ser Gly Lys Thr Pro Glu Ala Leu Ser Ala Gln Ala Thr Ala Leu			
1445	1450	1455	
atg acc tat ctg gac gag cga cct gat gtc tcc tcg ctg gat gtt ggg			4416

Met Thr Tyr Leu Asp Glu Arg Pro Asp Val Ser Ser Leu Asp Val Gly  
 1460 1465 1470

tac tcg ctg gcg ttg aca cgg tcg gcg ctg gat gag cga gcg gtg gtg 4464  
 Tyr Ser Leu Ala Leu Thr Arg Ser Ala Leu Asp Glu Arg Ala Val Val  
 1475 1480 1485

ctg ggg tcg gac cgt gaa acg ttg ttg tgc ggt gtg aaa gcg ctg tct 4512  
 Leu Gly Ser Asp Arg Glu Thr Leu Leu Cys Gly Val Lys Ala Leu Ser  
 1490 1495 1500

gcc ggt cat gag gct tct ggg ttg gtg acc gga tct gtg ggg gct ggg 4560  
 Ala Gly His Glu Ala Ser Gly Leu Val Thr Gly Ser Val Gly Ala Gly  
 1505 1510 1515 1520

ggc cgc atc ggg ttt gtg ttt tcc ggt cag ggt ggt cag tgg ctg ggg 4608  
 Gly Arg Ile Gly Phe Val Phe Ser Gly Gln Gly Gly Gln Trp Leu Gly  
 1525 1530 1535

atg ggc cgg ggg ctt tac cgg gct ttt ccg gtg ttc gct gct gcc ttt 4656  
 Met Gly Arg Gly Leu Tyr Arg Ala Phe Pro Val Phe Ala Ala Phe  
 1540 1545 1550

gac gaa gct tgt gcc gag ctg gat gcg cat ctg ggc cag gaa atc ggg 4704  
 Asp Glu Ala Cys Ala Glu Leu Asp Ala His Leu Gly Gln Glu Ile Gly  
 1555 1560 1565

gtt cgg gag gtg gtg tcc ggt tcg gat gcg cag ttg ctg gat cgg acg 4752  
 Val Arg Glu Val Val Ser Gly Ser Asp Ala Gln Leu Leu Asp Arg Thr  
 1570 1575 1580

ttg tgg gcg cag tcg ggt ttg ttc gcg ttg cag gtg ggc ttg ctg aag 4800  
 Leu Trp Ala Gln Ser Gly Leu Phe Ala Leu Gln Val Gly Leu Leu Lys  
 1585 1590 1595 1600

ttg ctg gat tcg tgg ggg gtt cgg ccg agt gtg gtg ttg ggg cat tcg 4848  
 Leu Leu Asp Ser Trp Gly Val Arg Pro Ser Val Val Leu Gly His Ser  
 1605 1610 1615

gtg ggc gag ttg gcg gcg gcg ttc gcg gcg ggt gtg gtg tcg ttg tcg 4896  
 Val Gly Glu Leu Ala Ala Ala Phe Ala Ala Gly Val Val Ser Leu Ser  
 1620 1625 1630

ggt gcg gct cgg ttg gtg gcg ggt cgt gcc cgg ttg atg cag gcg ttg 4944  
 Gly Ala Ala Arg Leu Val Ala Gly Arg Ala Arg Leu Met Gln Ala Leu  
 1635 1640 1645

ccg tct ggc ggt ggg atg ctg gcg gtg cct gct ggt gag gag ctg ttg 4992



His Val Arg Gly Gly Ala Val Asp Trp Arg Ser Phe Phe Ala Gly Thr  
 1845 1850 1855

agg gcg aag caa atc gag ctg ccc acc tac gcc ttc caa cga cag cgg 5616  
 Arg Ala Lys Gln Ile Glu Leu Pro Thr Tyr Ala Phe Gln Arg Gln Arg  
 1860 1865 1870

tac tgg ctg aac gcg ctg cgt gaa tct tcc gcc ggc gac atg ggc agg 5664  
 Tyr Trp Leu Asn Ala Leu Arg Glu Ser Ser Ala Gly Asp Met Gly Arg  
 1875 1880 1885

cgt gtc gaa gcg aag ttc tgg ggc gcc gtc gag cac gaa gat gtg gaa 5712  
 Arg Val Glu Ala Lys Phe Trp Gly Ala Val Glu His Glu Asp Val Glu  
 1890 1895 1900

tcg ctt gca cgc gta ttg ggc att gtg gac gac ggc gct gct gtg gat 5760  
 Ser Leu Ala Arg Val Leu Gly Ile Val Asp Asp Gly Ala Ala Val Asp  
 1905 1910 1915 1920

tcc ctg aga agc gcc ctt ccg gtg ttg gcc ggt tgg cag cga acc cgc 5808  
 Ser Leu Arg Ser Ala Leu Pro Val Leu Ala Gly Trp Gln Arg Thr Arg  
 1925 1930 1935

acc acc gag tcc att atg gat cag cgg tgt tac cga att ggc tgg cgg 5856  
 Thr Thr Glu Ser Ile Met Asp Gln Arg Cys Tyr Arg Ile Gly Trp Arg  
 1940 1945 1950

cag gta gcc gga ctc ccg ccg atg gga act gtt ttc ggt acc tgg ctg 5904  
 Gln Val Ala Gly Leu Pro Pro Met Gly Thr Val Phe Gly Thr Trp Leu  
 1955 1960 1965

gtc ttc gcg cct cat ggc tgg tcc agc gaa ccg gag gtg gtg gac tgc 5952  
 Val Phe Ala Pro His Gly Trp Ser Ser Glu Pro Glu Val Val Asp Cys  
 1970 1975 1980

gtt acg gca ctg ccg gca cgt ggt gcc tcg gtg gtg ttg gtg gaa gct 6000  
 Val Thr Ala Leu Arg Ala Arg Gly Ala Ser Val Val Leu Val Glu Ala  
 1985 1990 1995 2000

gat ccc gac ccg acc tcc ttc ggc gac ccg gta cga acc ctg tgt tcg 6048  
 Asp Pro Asp Pro Thr Ser Phe Gly Asp Arg Val Arg Thr Leu Cys Ser  
 2005 2010 2015

ggc ctt ccg gat ctt gtt ggc gtg ttg tca atg ttg tgc ttg gaa gaa 6096  
 Gly Leu Pro Asp Leu Val Gly Val Leu Ser Met Leu Cys Leu Glu Glu  
 2020 2025 2030

tcg gtc ctt ccg gga ttt tct gcg gtg tca ccg ggt ttt gcg ttg acc 6144

Ser Val Leu Pro G	Phe Ser Ala Val Ser Arg Gly P	Ala Leu Thr	
2035	2040	2045	
gtg gag ttg gtg cgg gtt ttg cgg gca gct ggt gcg act gcc cgg ttg			6192
Val Glu Leu Val Arg Val Leu Arg Ala Ala Gly Ala Thr Ala Arg Leu			
2050	2055	2060	
tgg ttg ctg acg tgt ggt ggc gtg tcg gtg gga gat gta ccg gtt cgt			6240
Trp Leu Leu Thr Cys Gly Gly Val Ser Val Gly Asp Val Pro Val Arg			
2065	2070	2075	2080
cca gcg cag gcc ctg gcg tgg ggg ttg ggg cgt gtt gtg ggg ttg gag			6288
Pro Ala Gln Ala Leu Ala Trp Gly Leu Gly Arg Val Val Gly Leu Glu			
2085	2090	2095	
cat ccg gac tgg tgg ggc ggc ttg atc gat att ccg gtc ttg ttc gac			6336
His Pro Asp Trp Trp Gly Gly Leu Ile Asp Ile Pro Val Leu Phe Asp			
2100	2105	2110	
gaa gac gct caa gag cgg ttg tcg att gtg ctg gca ggt ctc gat gag			6384
Glu Asp Ala Gln Glu Arg Leu Ser Ile Val Leu Ala Gly Leu Asp Glu			
2115	2120	2125	
gac gag gtc gcg atc cgt cct gac ggc atg ttc gcg cgt cgg ttg gta			6432
Asp Glu Val Ala Ile Arg Pro Asp Gly Met Phe Ala Arg Arg Leu Val			
2130	2135	2140	
cgc cac act gtc tca gct gat gtg aag aag gcg tgg cgc ccc agg gga			6480
Arg His Thr Val Ser Ala Asp Val Lys Lys Ala Trp Arg Pro Arg Gly			
2145	2150	2155	2160
tcg gtg ctg gtg acg ggc ggc acg ggt ggt ttg ggg gcg cac gtt gct			6528
Ser Val Leu Val Thr Gly Gly Thr Gly Gly Leu Gly Ala His Val Ala			
2165	2170	2175	
cgc tgg ctg gcc gac gcc gga gcc gaa cat gtg gcg atg gtg agt cga			6576
Arg Trp Leu Ala Asp Ala Gly Ala Glu His Val Ala Met Val Ser Arg			
2180	2185	2190	
cgc ggc gag cag gca ccg agt gct gag aag ttg cgg acg gaa ctg gag			6624
Arg Gly Glu Gln Ala Pro Ser Ala Glu Lys Leu Arg Thr Glu Leu Glu			
2195	2200	2205	
gat ctg ggt acc cgg gtg tcg atc gtg tca tgc gat gtg acc gat cgc			6672
Asp Leu Gly Thr Arg Val Ser Ile Val Ser Cys Asp Val Thr Asp Arg			
2210	2215	2220	
gag gcg ctc gcc gaa gtg ctg aaa gcc ctt ccg gct gaa aac ccg ttg			6720

Glu Ala Leu Ala Glu Val Leu Lys Ala Leu Pro Ala Glu Asn Pro Leu  
 2225                      2230                      2235                      2240

acc gcg gta gtg cat gcg gca ggc gtg atc gag act ggt gat gcg gcg      6768  
 Thr Ala Val Val His Ala Ala Gly Val Ile Glu Thr Gly Asp Ala Ala  
                                  2245                      2250                      2255

gca atg agc ctg gct gat ttc gat cac gtg ttg tcc gca aag gtg gcc      6816  
 Ala Met Ser Leu Ala Asp Phe Asp His Val Leu Ser Ala Lys Val Ala  
                                  2260                      2265                      2270

ggt gcc gcg aat ctg gat gcc ttg ttg gcc gat gtg gaa ttg gac gcg      6864  
 Gly Ala Ala Asn Leu Asp Ala Leu Leu Ala Asp Val Glu Leu Asp Ala  
                                  2275                      2280                      2285

ttc gtc ttg ttc tca tcg gtg tca gga gtt tgg ggc gct ggg gga cac      6912  
 Phe Val Leu Phe Ser Ser Val Ser Gly Val Trp Gly Ala Gly Gly His  
                                  2290                      2295                      2300

ggg gct tac gca gcg gcg aat gcc tat ctg gat gcg ctc gcg gaa cag      6960  
 Gly Ala Tyr Ala Ala Ala Asn Ala Tyr Leu Asp Ala Leu Ala Glu Gln  
 2305                      2310                      2315                      2320

cgt cgg tcg cga ggg ctg gtc gcg act gcg gtg gcc tgg ggg ccg tgg      7008  
 Arg Arg Ser Arg Gly Leu Val Ala Thr Ala Val Ala Trp Gly Pro Trp  
                                  2325                      2330                      2335

gcc ggc gag ggc atg gcc tcc gga gaa aca gga gac cag ctg cgc cga      7056  
 Ala Gly Glu Gly Met Ala Ser Gly Glu Thr Gly Asp Gln Leu Arg Arg  
                                  2340                      2345                      2350

tac ggc ctt tcc cca atg gct ccg cag cac gcc atc gcc gga atc cgg      7104  
 Tyr Gly Leu Ser Pro Met Ala Pro Gln His Ala Ile Ala Gly Ile Arg  
                                  2355                      2360                      2365

cag gcc gtg gaa cag gac gaa att tcc ctg gta gtg gcc gat gtc gat      7152  
 Gln Ala Val Glu Gln Asp Glu Ile Ser Leu Val Val Ala Asp Val Asp  
                                  2370                      2375                      2380

tgg gca cgt ttc agc gcg gga ttg ctg gcg gct agg ccg cgg ccg ctg      7200  
 Trp Ala Arg Phe Ser Ala Gly Leu Leu Ala Ala Arg Pro Arg Pro Leu  
 2385                      2390                      2395                      2400

ctg aac gaa ctg gcc gag gtc aag gaa ctc ctc gtc gat gcc cag ccc      7248  
 Leu Asn Glu Leu Ala Glu Val Lys Glu Leu Leu Val Asp Ala Gln Pro  
                                  2405                      2410                      2415

gag gcg gga gtc ctt gcc gac gcg tcg ttg gaa tgg cgg cag cga ttg      7296

Glu Ala Gly Val L	la Asp Ala Ser Leu Glu Trp A	ln Arg Leu	
2420	2425	2430	
tcc gcg gca ccg agg ccg aca cag gaa cag ctg atc ctg gag ctg gta			7344
Ser Ala Ala Pro Arg Pro Thr Gln Glu Gln Leu Ile Leu Glu Leu Val			
2435	2440	2445	
cgc ggc gaa acc gct ctg gtg ctg gga cac ccc ggg gca gcg gcc gtt			7392
Arg Gly Glu Thr Ala Leu Val Leu Gly His Pro Gly Ala Ala Ala Val			
2450	2455	2460	
gca tcg gaa cga gcc ttc aag gac agc gga ttc gac tcg cag gcc gcg			7440
Ala Ser Glu Arg Ala Phe Lys Asp Ser Gly Phe Asp Ser Gln Ala Ala			
2465	2470	2475 2480	
gtc gaa ctc cgc gtt cgg ctc aat cga gct acc ggc ctc cag ttg cca			7488
Val Glu Leu Arg Val Arg Leu Asn Arg Ala Thr Gly Leu Gln Leu Pro			
2485	2490	2495	
tcg aca att atc ttc agc cat ccc acg cct gcg gaa ctg gct gcg gag			7536
Ser Thr Ile Ile Phe Ser His Pro Thr Pro Ala Glu Leu Ala Ala Glu			
2500	2505	2510	
ctg cgg gcg agg ctt ctt ccc gag tcc gca gga gca ggc att ccc gag			7584
Leu Arg Ala Arg Leu Leu Pro Glu Ser Ala Gly Ala Gly Ile Pro Glu			
2515	2520	2525	
gag gac gag gcg cga atc aga gcg gca ctg acg tcg atc ccg ttc ccg			7632
Glu Asp Glu Ala Arg Ile Arg Ala Ala Leu Thr Ser Ile Pro Phe Pro			
2530	2535	2540	
gcc ttg cgc gag gca ggc ttg gtg agt ccg ctg ctc gca ctt gcc gga			7680
Ala Leu Arg Glu Ala Gly Leu Val Ser Pro Leu Leu Ala Leu Ala Gly			
2545	2550	2555 2560	
cac ccg gtc gac tcc ggt atc tcc tcg gac gat gcg gcc gcg acc tcg			7728
His Pro Val Asp Ser Gly Ile Ser Ser Asp Asp Ala Ala Ala Thr Ser			
2565	2570	2575	
atc gat gcg atg gat gta gcc ggc ctc gtc gaa gca gcg ctg ggc gaa			7776
Ile Asp Ala Met Asp Val Ala Gly Leu Val Glu Ala Ala Leu Gly Glu			
2580	2585	2590	
cgc gag tcc tga			7788
Arg Glu Ser			
2595			



<210> 42  
<211> 2595  
<212> PRT  
<213> Saccharopolyspora spinosa

<400> 42

Met Ser Glu Ala Gly Asn Leu Ile Ala Val Ile Gly Leu Ser Cys Arg  
1 5 10 15

Leu Pro Gln Ala Pro Asp Pro Ala Ser Phe Trp Arg Leu Leu Arg Thr  
20 25 30

Gly Thr Asp Ala Ile Thr Thr Val Pro Glu Gly Arg Trp Gly Asp Pro  
35 40 45

Leu Pro Gly Arg Asp Ala Pro Lys Gly Pro Glu Trp Gly Gly Phe Leu  
50 55 60

Ala Asp Val Asp Cys Phe Asp Pro Glu Phe Phe Gly Ile Ser Pro Arg  
65 70 75 80

Glu Ala Ala Ala Val Asp Pro Gln Gln Arg Leu Ala Leu Glu Leu Ala  
85 90 95

Trp Glu Ala Leu Glu Asp Ala Gly Ile Pro Ala Gly Glu Leu Arg Gly  
100 105 110

Thr Ala Ala Gly Val Phe Met Gly Ala Ile Ser Asp Asp Tyr Ala Ala  
115 120 125

Leu Leu Arg Glu Ser Pro Pro Glu Val Ala Ala Gln Tyr Arg Leu Thr  
130 135 140

Gly Thr His Arg Ser Leu Ile Ala Asn Arg Val Ser Tyr Val Leu Gly  
145 150 155 160

Leu Arg Gly Pro Ser Leu Thr Val Asp Ser Gly Gln Ser Ser Ser Leu  
165 170 175

Val Gly Val His Leu Ala Ser Glu Ser Leu Arg Arg Gly Glu Cys Thr  
180 185 190

Ile Ala Leu Ala Gly Gly Val Asn Leu Asn Leu Ala Ala Glu Ser Asn  
195 200 205

Ser Ala Leu Met Asp Phe Gly Ala Leu Ser Pro Asp Gly Arg Cys Phe  
210 215 220

Thr Phe Asp Val Ala Ala Asn Gly Tyr Val Arg Gly Gly Gly Gly  
225 230 235 240

Leu Val Val Leu Lys Lys Ala Asp Gln Ala His Ala Asp Gly Asp Arg  
245 250 255

Ile Tyr Cys Leu Ile Arg Gly Ser Ala Val Asn Asn Asp Gly Gly Gly  
260 265 270

Ala Gly Leu Thr Val Pro Ala Ala Asp Ala Gln Ala Glu Leu Leu Arg  
275 280 285

Gln Ala Tyr Arg Asn Ala Gly Val Asp Pro Ala Ala Val Gln Tyr Val  
290 295 300

Glu Leu His Gly Ser Ala Thr Arg Val Gly Asp Pro Val Glu Ala Ala  
305 310 315 320

Ala Leu Gly Ala Val Leu Gly Ala Ala Arg Arg Pro Gly Asp Glu Leu  
325 330 335

Arg Val Gly Ser Ala Lys Thr Asn Val Gly His Leu Glu Ala Ala Ala  
340 345 350

Gly Val Thr Gly Leu Leu Lys Thr Ala Leu Ser Ile Trp His Arg Glu  
355 360 365

Leu Pro Pro Ser Leu His Phe Thr Ala Pro Asn Pro Glu Ile Pro Leu  
370 375 380

Asp Glu Leu Asn Leu Arg Val Gln Arg Asp Leu Arg Pro Trp Pro Glu  
385 390 395 400

Ser Glu Gly Pro Leu Leu Ala Gly Val Ser Ala Phe Gly Met Gly Gly  
405 410 415

Thr Asn Cys His Leu Val Leu Ser Gly Thr Ser Arg Val Glu Arg Arg  
420 425 430

Arg Ser Gly Pro Ala Glu Ala Thr Met Pro Trp Val Leu Ser Ala Arg  
435 440 445

Thr Pro Val Ala Leu Arg Ala Gln Ala Ala Arg Leu His Thr His Leu  
450 455 460

Asn Thr Ala Gly Gln Ser Pro Leu Asp Val Ala Tyr Ser Leu Ala Thr  
465 470 475 480

Thr Arg Ser Ala Leu Pro His Arg Ala Ala Leu Val Ala Asp Asp Glu  
 485 490 495

Pro Lys Leu Leu Ala Gly Leu Lys Ala Leu Ala Asp Gly Asp Asp Ala  
 500 505 510

Pro Thr Leu Cys His Gly Ala Thr Ser Gly Glu Arg Ala Ala Val Phe  
 515 520 525

Val Phe Pro Gly Gln Gly Ser Gln Trp Ile Gly Met Gly Arg Gln Leu  
 530 535 540

Leu Glu Thr Ser Glu Val Phe Ala Ala Ser Met Ser Asp Cys Ala Asp  
 545 550 555 560

Ala Leu Ala Pro His Leu Asp Trp Ser Leu Leu Asp Val Leu Arg Asn  
 565 570 575

Ala Ala Gly Ala Ala His Leu Asp His Asp Asp Val Val Gln Pro Ala  
 580 585 590

Leu Phe Ala Ile Met Val Ser Leu Ala Glu Leu Trp Arg Ser Trp Gly  
 595 600 605

Val Arg Pro Val Ala Val Val Gly His Ser Gln Gly Glu Ile Ala Ala  
 610 615 620

Ala Cys Val Ala Gly Ala Leu Ser Val Arg Asp Ala Ala Arg Val Val  
 625 630 635 640

Ala Val Arg Ser Arg Leu Leu Thr Ala Leu Ala Gly Ser Gly Ala Met  
 645 650 655

Ala Ser Leu Gln His Pro Ala Glu Glu Val Arg Gln Ile Leu Leu Pro  
 660 665 670

Trp Arg Asp Arg Ile Gly Val Ala Gly Val Asn Gly Pro Ser Ser Thr  
 675 680 685

Leu Val Ser Gly Asp Arg Glu Ala Met Ala Glu Leu Leu Ala Glu Cys  
 690 695 700

Ala Asp Arg Glu Leu Arg Met Arg Arg Ile Pro Val Glu Tyr Ala Ser  
 705 710 715 720

His Ser Pro His Ile Glu Val Val Arg Asp Glu Leu Leu Gly Leu Leu  
 725 730 735

Ala Pro Val Glu Phe Arg Thr Gly Ser Ile Pro Ile Tyr Ser Thr Thr  
740 745 750

Thr Gly Asp Leu Leu Asp Arg Pro Met Asp Ala Asp Tyr Trp Tyr Arg  
755 760 765

Asn Leu Arg Gln Pro Val Leu Phe Glu Ala Ala Val Glu Ala Leu Leu  
770 775 780

Lys Arg Gly Tyr Asp Ala Phe Ile Glu Ile Ser Pro His Pro Val Leu  
785 790 795 800

Thr Ala Asn Ile Gln Glu Thr Ala Val Arg Ala Gly Arg Glu Val Val  
805 810 815

Ala Leu Gly Thr Leu Arg Arg Gly Glu Gly Gly Met Arg Gln Ala Leu  
820 825 830

Thr Ser Leu Ala Arg Ala His Val His Gly Val Ala Ala Asp Trp His  
835 840 845

Ala Val Phe Ala Gly Thr Gly Ala Gln Arg Val Asp Leu Pro Thr Tyr  
850 855 860

Ala Phe Gln Arg Gln Arg Tyr Trp Leu Asp Ala Lys Leu Pro Asp Val  
865 870 875 880

Ala Met Pro Glu Ser Asp Val Ser Thr Ala Leu Arg Glu Lys Leu Arg  
885 890 895

Ser Ser Pro Arg Ala Asp Val Asp Ser Thr Thr Leu Thr Met Ile Arg  
900 905 910

Ala Gln Ala Ala Val Val Leu Gly His Ser Asp Pro Lys Glu Val Asp  
915 920 925

Pro Asp Arg Thr Phe Lys Asp Leu Gly Phe Asp Ser Ser Met Val Val  
930 935 940

Glu Leu Cys Asp Arg Leu Asn Ala Ala Thr Gly Leu Arg Leu Ala Pro  
945 950 955 960

Ser Val Val Phe Asp Cys Pro Thr Pro Asp Lys Leu Ala Arg Gln Val  
965 970 975

Arg Thr Leu Leu Leu Gly Glu Pro Ala Pro Met Thr Ser His Arg Pro  
980 985 990

Asp Ser Asp Ala Asp Glu Pro Ile Ala Val Ile Gly Met Gly Cys Arg  
 995 1000 1005

Phe Pro Gly Gly Val Ser Ser Pro Glu Glu Leu Trp Gln Leu Val Ala  
 1010 1015 1020

Ala Gly Arg Asp Val Val Ser Glu Phe Pro Ala Asp Arg Gly Trp Asp  
 025 1030 1035 1040

Leu Glu Arg Ala Gly Thr Ser His Val Arg Ala Gly Gly Phe Leu His  
 1045 1050 1055

Gly Ala Pro Asp Phe Asp Pro Gly Phe Phe Arg Ile Ser Pro Arg Glu  
 1060 1065 1070

Ala Leu Ala Met Asp Pro Gln Gln Arg Leu Leu Leu Glu Ile Ala Trp  
 1075 1080 1085

Glu Ala Val Glu Arg Gly Gly Ile Asn Pro Gln His Leu His Gly Ser  
 1090 1095 1100

Gln Thr Gly Val Phe Val Gly Ala Thr Ser Leu Asp Tyr Gly Pro Arg  
 105 1110 1115 1120

Leu His Glu Ala Ser Glu Glu Ala Ala Gly Tyr Val Leu Thr Gly Ser  
 1125 1130 1135

Thr Thr Ser Val Ala Ser Gly Arg Val Ala Tyr Ser Phe Gly Phe Glu  
 1140 1145 1150

Gly Pro Ala Val Thr Val Asp Thr Ala Cys Ser Ser Ser Leu Val Ala  
 1155 1160 1165

Leu His Leu Ala Cys Gln Ser Leu Arg Ser Gly Glu Cys Asp Leu Ala  
 1170 1175 1180

Leu Ala Gly Gly Val Thr Val Met Ala Thr Pro Gly Met Phe Val Glu  
 185 1190 1195 1200

Phe Ser Arg Gln Arg Gly Leu Ala Pro Asp Gly Arg Cys Lys Ser Phe  
 1205 1210 1215

Ala Glu Ala Ala Asp Gly Thr Gly Trp Ser Glu Gly Ala Gly Leu Val  
 1220 1225 1230

Leu Leu Glu Arg Leu Ser Asp Ala Arg Arg Asn Gly His Glu Val Leu  
 1235 1240 1245

Ala Val Val Arg Ser Ala Val Asn Gln Asp Gly Ser Asn Gly  
1250 1255 1260

Leu Thr Ala Pro Asn Gly Ser Ser Gln Gln Arg Val Ile Ala Gln Ala  
265 1270 1275 1280

Leu Ala Ser Ala Gly Leu Ser Val Ser Asp Val Asp Ala Val Glu Ala  
1285 1290 1295

His Gly Thr Gly Thr Arg Leu Gly Asp Pro Ile Glu Ala Gln Ala Leu  
1300 1305 1310

Ile Ala Thr Tyr Gly Gln Gly Arg Leu Pro Glu Arg Pro Leu Trp Leu  
1315 1320 1325

Gly Ser Met Lys Ser Asn Ile Gly His Ala Gln Ala Ala Ala Gly Ile  
1330 1335 1340

Ala Gly Val Met Lys Met Val Met Ala Met Arg His Gly Gln Leu Pro  
345 1350 1355 1360

Arg Thr Leu His Val Asp Glu Pro Thr Ser Gly Val Asp Trp Ser Ala  
1365 1370 1375

Gly Thr Val Gln Leu Leu Thr Glu Asn Thr Pro Trp Pro Gly Ser Gly  
1380 1385 1390

Arg Val Arg Arg Val Gly Val Ser Ser Phe Gly Ile Ser Gly Thr Asn  
1395 1400 1405

Ala His Val Ile Leu Glu Gln Pro Pro Gly Val Pro Ser Gln Ser Ala  
1410 1415 1420

Gly Pro Gly Ser Gly Ser Val Val Asp Val Pro Val Val Pro Trp Met  
425 1430 1435 1440

Val Ser Gly Lys Thr Pro Glu Ala Leu Ser Ala Gln Ala Thr Ala Leu  
1445 1450 1455

Met Thr Tyr Leu Asp Glu Arg Pro Asp Val Ser Ser Leu Asp Val Gly  
1460 1465 1470

Tyr Ser Leu Ala Leu Thr Arg Ser Ala Leu Asp Glu Arg Ala Val Val  
1475 1480 1485

Leu Gly Ser Asp Arg Glu Thr Leu Leu Cys Gly Val Lys Ala Leu Ser  
1490 1495 1500

Ala Gly His Glu Ala Ser Gly Leu Val Thr Gly Ser Val Gly Ala Gly  
505 1510 1515 1520

Gly Arg Ile Gly Phe Val Phe Ser Gly Gln Gly Gly Gln Trp Leu Gly  
1525 1530 1535

Met Gly Arg Gly Leu Tyr Arg Ala Phe Pro Val Phe Ala Ala Ala Phe  
1540 1545 1550

Asp Glu Ala Cys Ala Glu Leu Asp Ala His Leu Gly Gln Glu Ile Gly  
1555 1560 1565

Val Arg Glu Val Val Ser Gly Ser Asp Ala Gln Leu Leu Asp Arg Thr  
1570 1575 1580

Leu Trp Ala Gln Ser Gly Leu Phe Ala Leu Gln Val Gly Leu Leu Lys  
585 1590 1595 1600

Leu Leu Asp Ser Trp Gly Val Arg Pro Ser Val Val Leu Gly His Ser  
1605 1610 1615

Val Gly Glu Leu Ala Ala Ala Phe Ala Ala Gly Val Val Ser Leu Ser  
1620 1625 1630

Gly Ala Ala Arg Leu Val Ala Gly Arg Ala Arg Leu Met Gln Ala Leu  
1635 1640 1645

Pro Ser Gly Gly Gly Met Leu Ala Val Pro Ala Gly Glu Glu Leu Leu  
1650 1655 1660

Trp Ser Leu Leu Ala Asp Gln Gly Asp Arg Val Gly Ile Ala Ala Val  
665 1670 1675 1680

Asn Ala Ala Gly Ser Val Val Leu Ser Gly Asp Arg Asp Val Leu Asp  
1685 1690 1695

Asp Leu Ala Gly Arg Leu Asp Gly Gln Gly Ile Arg Ser Arg Trp Leu  
1700 1705 1710

Arg Val Ser His Ala Phe His Ser Tyr Arg Met Asp Pro Met Leu Ala  
1715 1720 1725

Glu Phe Ala Glu Leu Ala Arg Thr Val Asp Tyr Arg Arg Cys Glu Val  
1730 1735 1740

Pro Ile Val Ser Thr Leu Thr Gly Asp Leu Asp Asp Ala Gly Arg Met  
745 1750 1755 1760

Ser Gly Pro Asp Trp Val Arg Gln Val Arg Glu Val Arg Phe  
 1765 1770 1775

Ala Asp Gly Val Gln Ala Leu Val Glu His Asp Val Ala Thr Val Val  
 1780 1785 1790

Glu Leu Gly Pro Asp Gly Ala Leu Ser Ala Leu Ile Gln Glu Cys Val  
 1795 1800 1805

Ala Ala Ser Asp His Ala Gly Arg Leu Ser Ala Val Pro Ala Met Arg  
 1810 1815 1820

Arg Asn Gln Asp Glu Ala Gln Lys Val Met Thr Ala Leu Ala His Val  
 825 1830 1835 1840

His Val Arg Gly Gly Ala Val Asp Trp Arg Ser Phe Phe Ala Gly Thr  
 1845 1850 1855

Arg Ala Lys Gln Ile Glu Leu Pro Thr Tyr Ala Phe Gln Arg Gln Arg  
 1860 1865 1870

Tyr Trp Leu Asn Ala Leu Arg Glu Ser Ser Ala Gly Asp Met Gly Arg  
 1875 1880 1885

Arg Val Glu Ala Lys Phe Trp Gly Ala Val Glu His Glu Asp Val Glu  
 1890 1895 1900

Ser Leu Ala Arg Val Leu Gly Ile Val Asp Asp Gly Ala Ala Val Asp  
 905 1910 1915 1920

Ser Leu Arg Ser Ala Leu Pro Val Leu Ala Gly Trp Gln Arg Thr Arg  
 1925 1930 1935

Thr Thr Glu Ser Ile Met Asp Gln Arg Cys Tyr Arg Ile Gly Trp Arg  
 1940 1945 1950

Gln Val Ala Gly Leu Pro Pro Met Gly Thr Val Phe Gly Thr Trp Leu  
 1955 1960 1965

Val Phe Ala Pro His Gly Trp Ser Ser Glu Pro Glu Val Val Asp Cys  
 1970 1975 1980

Val Thr Ala Leu Arg Ala Arg Gly Ala Ser Val Val Leu Val Glu Ala  
 985 1990 1995 2000

Asp Pro Asp Pro Thr Ser Phe Gly Asp Arg Val Arg Thr Leu Cys Ser  
 2005 2010 2015



Gly Leu Pro Asp Leu Val Gly Val Leu Ser Met Leu Cys Leu Glu Glu  
 2020 2025 2030

Ser Val Leu Pro Gly Phe Ser Ala Val Ser Arg Gly Phe Ala Leu Thr  
 2035 2040 2045

Val Glu Leu Val Arg Val Leu Arg Ala Ala Gly Ala Thr Ala Arg Leu  
 2050 2055 2060

Trp Leu Leu Thr Cys Gly Gly Val Ser Val Gly Asp Val Pro Val Arg  
 065 2070 2075 2080

Pro Ala Gln Ala Leu Ala Trp Gly Leu Gly Arg Val Val Gly Leu Glu  
 2085 2090 2095

His Pro Asp Trp Trp Gly Gly Leu Ile Asp Ile Pro Val Leu Phe Asp  
 2100 2105 2110

Glu Asp Ala Gln Glu Arg Leu Ser Ile Val Leu Ala Gly Leu Asp Glu  
 2115 2120 2125

Asp Glu Val Ala Ile Arg Pro Asp Gly Met Phe Ala Arg Arg Leu Val  
 2130 2135 2140

Arg His Thr Val Ser Ala Asp Val Lys Lys Ala Trp Arg Pro Arg Gly  
 145 2150 2155 2160

Ser Val Leu Val Thr Gly Gly Thr Gly Gly Leu Gly Ala His Val Ala  
 2165 2170 2175

Arg Trp Leu Ala Asp Ala Gly Ala Glu His Val Ala Met Val Ser Arg  
 2180 2185 2190

Arg Gly Glu Gln Ala Pro Ser Ala Glu Lys Leu Arg Thr Glu Leu Glu  
 2195 2200 2205

Asp Leu Gly Thr Arg Val Ser Ile Val Ser Cys Asp Val Thr Asp Arg  
 2210 2215 2220

Glu Ala Leu Ala Glu Val Leu Lys Ala Leu Pro Ala Glu Asn Pro Leu  
 225 2230 2235 2240

Thr Ala Val Val His Ala Ala Gly Val Ile Glu Thr Gly Asp Ala Ala  
 2245 2250 2255

Ala Met Ser Leu Ala Asp Phe Asp His Val Leu Ser Ala Lys Val Ala  
 2260 2265 2270

Gly Ala Ala Asn Leu Asp Ala Leu Leu Ala Asp Val Glu Leu Asp Ala  
 2275 2280 2285

Phe Val Leu Phe Ser Ser Val Ser Gly Val Trp Gly Ala Gly Gly His  
 2290 2295 2300

Gly Ala Tyr Ala Ala Ala Asn Ala Tyr Leu Asp Ala Leu Ala Glu Gln  
 305 2310 2315 2320

Arg Arg Ser Arg Gly Leu Val Ala Thr Ala Val Ala Trp Gly Pro Trp  
 2325 2330 2335

Ala Gly Glu Gly Met Ala Ser Gly Glu Thr Gly Asp Gln Leu Arg Arg  
 2340 2345 2350

Tyr Gly Leu Ser Pro Met Ala Pro Gln His Ala Ile Ala Gly Ile Arg  
 2355 2360 2365

Gln Ala Val Glu Gln Asp Glu Ile Ser Leu Val Val Ala Asp Val Asp  
 2370 2375 2380

Trp Ala Arg Phe Ser Ala Gly Leu Leu Ala Ala Arg Pro Arg Pro Leu  
 385 2390 2395 2400

Leu Asn Glu Leu Ala Glu Val Lys Glu Leu Leu Val Asp Ala Gln Pro  
 2405 2410 2415

Glu Ala Gly Val Leu Ala Asp Ala Ser Leu Glu Trp Arg Gln Arg Leu  
 2420 2425 2430

Ser Ala Ala Pro Arg Pro Thr Gln Glu Gln Leu Ile Leu Glu Leu Val  
 2435 2440 2445

Arg Gly Glu Thr Ala Leu Val Leu Gly His Pro Gly Ala Ala Ala Val  
 2450 2455 2460

Ala Ser Glu Arg Ala Phe Lys Asp Ser Gly Phe Asp Ser Gln Ala Ala  
 465 2470 2475 2480

Val Glu Leu Arg Val Arg Leu Asn Arg Ala Thr Gly Leu Gln Leu Pro  
 2485 2490 2495

Ser Thr Ile Ile Phe Ser His Pro Thr Pro Ala Glu Leu Ala Ala Glu  
 2500 2505 2510

Leu Arg Ala Arg Leu Leu Pro Glu Ser Ala Gly Ala Gly Ile Pro Glu  
 2515 2520 2525

Glu Asp Glu Ala Arg Ile Arg Ala Ala Leu Thr Ser Ile Pro Phe Pro  
 2530 2535 2540

Ala Leu Arg Glu Ala Gly Leu Val Ser Pro Leu Leu Ala Leu Ala Gly  
 545 2550 2555 2560

His Pro Val Asp Ser Gly Ile Ser Ser Asp Asp Ala Ala Ala Thr Ser  
 2565 2570 2575

Ile Asp Ala Met Asp Val Ala Gly Leu Val Glu Ala Ala Leu Gly Glu  
 2580 2585 2590

Arg Glu Ser  
 2595

<210> 43  
 <211> 6459  
 <212> DNA  
 <213> Saccharopolyspora spinosa

<220>  
 <221> CDS  
 <222> (1)..(6456)  
 <223> ORF19; Polyketidsynthase

<400> 43  
 atg acg gtg acc acc agt tac gaa gaa gtt gtc gag gca ctg cga gca 48  
 Met Thr Val Thr Thr Ser Tyr Glu Glu Val Val Glu Ala Leu Arg Ala  
 1 5 10 15

tcg ctc aag gag aac gaa cgc ctc cgg cgc ggc agg gat cgg ttc tcc 96  
 Ser Leu Lys Glu Asn Glu Arg Leu Arg Arg Gly Arg Asp Arg Phe Ser  
 20 25 30

gcg gag aag gac gat ccc atc gcg atc gtg gcg atg agt tgt cgt tat 144  
 Ala Glu Lys Asp Asp Pro Ile Ala Ile Val Ala Met Ser Cys Arg Tyr  
 35 40 45

ccc ggt cag gtc tcc tcg ccg gag gac ctg tgg caa ctg gct gcc ggc 192  
 Pro Gly Gln Val Ser Ser Pro Glu Asp Leu Trp Gln Leu Ala Ala Gly  
 50 55 60

ggt gtg gac gcg atc tcc gaa gtt ccg ggg gat cgc gga tgg gac ctg 240  
 Gly Val Asp Ala Ile Ser Glu Val Pro Gly Asp Arg Gly Trp Asp Leu  
 65 70 75 80

gat ggc gtg ttc ggc ccg gac tcc gat cgt cct ggc acg tcg tat gcc	288
Asp Gly Val Phe Val Pro Asp Ser Asp Arg Pro Gly Thr Ser Tyr Ala	
85 90 95	
tgc gcg ggc ggt ttt ctt cag ggc gtg tcg gag ttc gac gcg ggt ttc	336
Cys Ala Gly Gly Phe Leu Gln Gly Val Ser Glu Phe Asp Ala Gly Phe	
100 105 110	
ttc ggg att tcg ccg cgt gag gcg ctg gcg atg gat ccg cag cag cgg	384
Phe Gly Ile Ser Pro Arg Glu Ala Leu Ala Met Asp Pro Gln Gln Arg	
115 120 125	
ttg ctg ctg gaa gtc gcg tgg gag gtc ttc gag cgg gct ggg ctg gag	432
Leu Leu Leu Glu Val Ala Trp Glu Val Phe Glu Arg Ala Gly Leu Glu	
130 135 140	
cag cgg tcg aca cgc ggt tcc cgc gtt ggc gtg ttc gtc ggc acc aat	480
Gln Arg Ser Thr Arg Gly Ser Arg Val Gly Val Phe Val Gly Thr Asn	
145 150 155 160	
ggc cag gac tac gcg tcg tgg ttg cgg acg ccg ccg cct gcg gtg gca	528
Gly Gln Asp Tyr Ala Ser Trp Leu Arg Thr Pro Pro Pro Ala Val Ala	
165 170 175	
ggt cat gtg ctg acg ggc ggt gcg gca gcg gtt ctt tcg ggc cgg gtt	576
Gly His Val Leu Thr Gly Gly Ala Ala Ala Val Leu Ser Gly Arg Val	
180 185 190	
gcg tat tcg ttc ggg ttc gag ggt cct gcg gtg acg gtg gat acg gcg	624
Ala Tyr Ser Phe Gly Phe Glu Gly Pro Ala Val Thr Val Asp Thr Ala	
195 200 205	
tgt tcg tcg tcg ttg gtg gcg ttg cac ctg gcg ggg caa gca ctg cgg	672
Cys Ser Ser Ser Leu Val Ala Leu His Leu Ala Gly Gln Ala Leu Arg	
210 215 220	
gcc ggt gag tgc gac ctt gcc ctt gcc ggt ggc gtc acg gtg atg tcg	720
Ala Gly Glu Cys Asp Leu Ala Leu Ala Gly Gly Val Thr Val Met Ser	
225 230 235 240	
acg ccg aag gtg ttc ctg gag ttc tcc cgc caa cgg ggt ctc gcg ccg	768
Thr Pro Lys Val Phe Leu Glu Phe Ser Arg Gln Arg Gly Leu Ala Pro	
245 250 255	
gat ggg cgg tgc aag tcg ttc gcg gcg ggt gcg gat ggc act gga tgg	816
Asp Gly Arg Cys Lys Ser Phe Ala Ala Gly Ala Asp Gly Thr Gly Trp	
260 265 270	

ggt gag ggt gcc gga ctg ttg ttg ctg gag cgg ttg tcg gat gcc cgg	864
Gly Glu Gly Ala Gly Leu Leu Leu Leu Glu Arg Leu Ser Asp Ala Arg	
275 280 285	
cgg aat ggg cat gag gtg ctg gcg gtt gtt cgt ggt agt gcg gtg aat	912
Arg Asn Gly His Glu Val Leu Ala Val Val Arg Gly Ser Ala Val Asn	
290 295 300	
cag gac ggt gcg tcg aat ggt ttg acc gcg ccg aat ggt tcg tcg cag	960
Gln Asp Gly Ala Ser Asn Gly Leu Thr Ala Pro Asn Gly Ser Ser Gln	
305 310 315 320	
cag cgg gtg att acc cag gcg ttg gcg agt gcg ggg ttg tcg gtg tcc	1008
Gln Arg Val Ile Thr Gln Ala Leu Ala Ser Ala Gly Leu Ser Val Ser	
325 330 335	
gat gtg gat gct gtg gag gcg cat ggg acg ggc acg cgg ctt ggt gat	1056
asp Val Asp Ala Val Glu Ala His Gly Thr Gly Thr Arg Leu Gly Asp	
340 345 350	
ccg atc gag gcg cag gcg ctg atc gcc acc tac ggc cgt gat cgt gat	1104
Pro Ile Glu Ala Gln Ala Leu Ile Ala Thr Tyr Gly Arg Asp Arg Asp	
355 360 365	
cct ggc cgg ccg ttg tgg ttg ggg tcg gtc aag tcg aac atc ggt cat	1152
Pro Gly Arg Pro Leu Trp Leu Gly Ser Val Lys Ser Asn Ile Gly His	
370 375 380	
acg caa gcg gcg gcg ggt gtg gct ggt gtg atc aag atg gtg atg gcg	1200
Thr Gln Ala Ala Ala Gly Val Ala Gly Val Ile Lys Met Val Met Ala	
385 390 395 400	
atg cgg cac ggg cag ctg cca cgc acg ttg cac gtg gaa tcg ccg tcg	1248
Met Arg His Gly Gln Leu Pro Arg Thr Leu His Val Glu Ser Pro Ser	
405 410 415	
ccg gag gtg gat tgg tcg gcg ggg acg gtt caa ctc ctt acg gag aac	1296
Pro Glu Val Asp Trp Ser Ala Gly Thr Val Gln Leu Leu Thr Glu Asn	
420 425 430	
acg ccc tgg ccc agg agt ggt cgt gtt cgt cgg gtg ggg gtg tcg tcg	1344
Thr Pro Trp Pro Arg Ser Gly Arg Val Arg Arg Val Gly Val Ser Ser	
435 440 445	
ttc ggg atc agt ggt act aac gcg cac gtc atc ctc gaa cag ccc ccg	1392
Phe Gly Ile Ser Gly Thr Asn Ala His Val Ile Leu Glu Gln Pro Pro	
450 455 460	

gga gtg ccg agt cct gcg ggg ccg ggt tcg ggt ttc gtc gtg gat	1440
Gly Val Pro Ser Gln Ser Ala Gly Pro Gly Ser Gly Ser Val Val Asp	
465 470 475 480	
gtt ccg gtg gtg ccg tgg atg gtg tcg ggc aaa aca ccc gaa gcg cta	1488
Val Pro Val Val Pro Trp Met Val Ser Gly Lys Thr Pro Glu Ala Leu	
485 490 495	
tcc gcg cag gca acg gcg ttg atg acc tat ctg gac gag cga cct gat	1536
Ser Ala Gln Ala Thr Ala Leu Met Thr Tyr Leu Asp Glu Arg Pro Asp	
500 505 510	
gtc tcc tcg ctg gat gtt ggg tac tcg ctg gcg ttg aca cgg tcg gcg	1584
Val Ser Ser Leu Asp Val Gly Tyr Ser Leu Ala Leu Thr Arg Ser Ala	
515 520 525	
ctg gat gag cga gcg gtg gtg ctg ggg tcg gac cgt gaa acg ttg ttg	1632
Leu Asp Glu Arg Ala Val Val Leu Gly Ser Asp Arg Glu Thr Leu Leu	
530 535 540	
tgc ggt gtg aaa gcg ctg tct gcc ggt cat gag gct tct ggg ttg gtg	1680
Cys Gly Val Lys Ala Leu Ser Ala Gly His Glu Ala Ser Gly Leu Val	
545 550 555 560	
acc gga tct gtg ggg gct ggg ggc cgc atc ggg ttt gtg ttt tcc ggt	1728
Thr Gly Ser Val Gly Ala Gly Gly Arg Ile Gly Phe Val Phe Ser Gly	
565 570 575	
cag ggt ggt cag tgg ctg ggg atg ggc cgg ggg ctt tac cgg gct ttt	1776
Gln Gly Gly Gln Trp Leu Gly Met Gly Arg Gly Leu Tyr Arg Ala Phe	
580 585 590	
ccg gtg ttc gct gct gcc ttt gac gaa gct tgt gcc gag ctg gat gca	1824
Pro Val Phe Ala Ala Ala Phe Asp Glu Ala Cys Ala Glu Leu Asp Ala	
595 600 605	
cat ctg ggc cag gaa atc ggg gtt cgg gag gtg gtg tcc ggt tcg gat	1872
His Leu Gly Gln Glu Ile Gly Val Arg Glu Val Val Ser Gly Ser Asp	
610 615 620	
gcg cag ttg ctg gat cgg acg ttg tgg gcg cag tcg ggt ttg ttc gcg	1920
Ala Gln Leu Leu Asp Arg Thr Leu Trp Ala Gln Ser Gly Leu Phe Ala	
625 630 635 640	
ttg cag gtg ggc ttg ctg aag ttg ctg gat tcg tgg ggg gtt cgg ccg	1968
Leu Gln Val Gly Leu Leu Lys Leu Leu Asp Ser Trp Gly Val Arg Pro	
645 650 655	

agt gtg gtg ttg ggg cat tcg gtg ggc gag ttg gcg gcg gcg ttc gcg 2016  
 Ser Val Val Leu Gly His Ser Val Gly Glu Leu Ala Ala Ala Phe Ala  
 660 665 670

gcg ggt gtg gtg tcg ttg tcg ggt gcg gct cgg ttg gtg gcg ggt cgt 2064  
 Ala Gly Val Val Ser Leu Ser Gly Ala Ala Arg Leu Val Ala Gly Arg  
 675 680 685

gcc cgg ttg atg cag gcg ttg ccg tct ggc ggt ggg atg ctg gcg gtg 2112  
 Ala Arg Leu Met Gln Ala Leu Pro Ser Gly Gly Gly Met Leu Ala Val  
 690 695 700

cct gct ggt gag gag ctg ttg tgg tcg ttg ttg gcc gat cag ggt gat 2160  
 Pro Ala Gly Glu Glu Leu Leu Trp Ser Leu Leu Ala Asp Gln Gly Asp  
 705 710 715 720

cgt gtg ggg atc gcc gcg gtc aac gct gcg ggg tcg gtg gtg ctc tct 2208  
 Arg Val Gly Ile Ala Ala Val Asn Ala Ala Gly Ser Val Val Leu Ser  
 725 730 735

ggt gat cgg gat gtg ctc gat gac ctt gcc ggt cgg ctg gac ggg caa 2256  
 Gly Asp Arg Asp Val Leu Asp Asp Leu Ala Gly Arg Leu Asp Gly Gln  
 740 745 750

ggg atc cgg tcg agg tgg ttg cgg gtg tcg cat gcg ttt cat tcg tat 2304  
 Gly Ile Arg Ser Arg Trp Leu Arg Val Ser His Ala Phe His Ser Tyr  
 755 760 765

cgg atg gat ccg atg ctg gcg gag ttc gcc gaa ttg gca cga acc gtg 2352  
 Arg Met Asp Pro Met Leu Ala Glu Phe Ala Glu Leu Ala Arg Thr Val  
 770 775 780

gat tac cgg cgt tgt gaa gtg ccg atc gtg tcg acc ttg acc gga gac 2400  
 Asp Tyr Arg Arg Cys Glu Val Pro Ile Val Ser Thr Leu Thr Gly Asp  
 785 790 795 800

ctc gat gac gct ggc agg atg agc ggg ccc gac tac tgg gtg cgt cag 2448  
 Leu Asp Asp Ala Gly Arg Met Ser Gly Pro Asp Tyr Trp Val Arg Gln  
 805 810 815

gtg cga gag ccg gtc cgc ttc gcc gac ggt gtc cag gcg ctg gtc gag 2496  
 Val Arg Glu Pro Val Arg Phe Ala Asp Gly Val Gln Ala Leu Val Glu  
 820 825 830

cac gat gtg gcc act gtt gtc gag ctc ggt ccg gac ggg gcg ttg tcg 2544  
 His Asp Val Ala Thr Val Val Glu Leu Gly Pro Asp Gly Ala Leu Ser  
 835 840 845

gcg ctg atc cag ggt gtc gcc gca tcc gat cac ggg cgg ctg 2592  
Ala Leu Ile Gln Glu Cys Val Ala Ala Ser Asp His Ala Gly Arg Leu  
850 855 860

agc gcg gtc ccg gcg atg cgc agg aac cag gac gag gcg cag aag gtg 2640  
Ser Ala Val Pro Ala Met Arg Arg Asn Gln Asp Glu Ala Gln Lys Val  
865 870 875 880

atg acg gcc ctg gca cac gtc cac gta cgt ggt ggt gcg gtg gac tgg 2688  
Met Thr Ala Leu Ala His Val His Val Arg Gly Gly Ala Val Asp Trp  
885 890 895

cgg tcg ttc ttc gcc ggt acg gga gcg aaa caa atc gag ctg ccc acc 2736  
Arg Ser Phe Phe Ala Gly Thr Gly Ala Lys Gln Ile Glu Leu Pro Thr  
900 905 910

tac gcc ttc caa cga cag cgg tac tgg ctg gtg cca tcg gat tcc ggt 2784  
Tyr Ala Phe Gln Arg Gln Arg Tyr Trp Leu Val Pro Ser Asp Ser Gly  
915 920 925

gat gtg aca ggt gcc ggt ctg gcc ggg gcg gag cat ccg ctg ttg ggt 2832  
Asp Val Thr Gly Ala Gly Leu Ala Gly Ala Glu His Pro Leu Leu Gly  
930 935 940

gct gtg gtg ccg gtc gcg ggt ggt gac gag gtg ttg ctg acc ggc agg 2880  
Ala Val Val Pro Val Ala Gly Gly Asp Glu Val Leu Leu Thr Gly Arg  
945 950 955 960

att tcg gtg cgg acg cat ccg tgg ctg gcc gaa cac cgg gtg ctg ggt 2928  
Ile Ser Val Arg Thr His Pro Trp Leu Ala Glu His Arg Val Leu Gly  
965 970 975

gaa gtg atc gtt gcg ggc acc gcg ttg ctg gag atc gcc ttg cac gcg 2976  
Glu Val Ile Val Ala Gly Thr Ala Leu Leu Glu Ile Ala Leu His Ala  
980 985 990

ggg gaa cgt ctt ggt tgt gaa cgg gtg gaa gag ctc acc ctg gaa gca 3024  
Gly Glu Arg Leu Gly Cys Glu Arg Val Glu Glu Leu Thr Leu Glu Ala  
995 1000 1005

ccg ctg gtc ctg ccg gag cgc ggg gcg atc cag gtt cag ctg cga gtg 3072  
Pro Leu Val Leu Pro Glu Arg Gly Ala Ile Gln Val Gln Leu Arg Val  
1010 1015 1020

ggc gcg ccc gag aat tcc gga cgc agg ccg atg gcg ctg tat tca cgc 3120  
Gly Ala Pro Glu Asn Ser Gly Arg Arg Pro Met Ala Leu Tyr Ser Arg  
1025 1030 1035 1040



ccc gaa ggg gcg gcg gag cat gac tgg acg cgg cac gcc acg ggc cgg 3168  
Pro Glu Gly Ala Ala Glu His Asp Trp Thr Arg His Ala Thr Gly Arg  
1045 1050 1055

ttg gcg cca ggc cgc ggc gag gcg gct gga gac ctg gcc gac tgg ccg 3216  
Leu Ala Pro Gly Arg Gly Glu Ala Ala Gly Asp Leu Ala Asp Trp Pro  
1060 1065 1070

gct cct ggc gcg ctg ccg gtc gac ctc gac gaa ttc tat cgg gac ctc 3264  
Ala Pro Gly Ala Leu Pro Val Asp Leu Asp Glu Phe Tyr Arg Asp Leu  
1075 1080 1085

gca gag ctt ggg ctg gag tac ggc ccg atc ttc caa ggg ctc aag gcg 3312  
Ala Glu Leu Gly Leu Glu Tyr Gly Pro Ile Phe Gln Gly Leu Lys Ala  
1090 1095 1100

gcc tgg cgg caa ggg gac gag gtg tac gcc gaa gcc gcg ctg ccg gga 3360  
Ala Trp Arg Gln Gly Asp Glu Val Tyr Ala Glu Ala Ala Leu Pro Gly  
1105 1110 1115 1120

acg gaa gat tct ggt ttc ggg gtg cat ccg gca ctg ctg gac gcg gct 3408  
Thr Glu Asp Ser Gly Phe Gly Val His Pro Ala Leu Leu Asp Ala Ala  
1125 1130 1135

ctg cac gca acg gct gtc cga gac atg gat gac gca cgc ttg ccg ttc 3456  
Leu His Ala Thr Ala Val Arg Asp Met Asp Asp Ala Arg Leu Pro Phe  
1140 1145 1150

cag tgg gaa ggt gtg tcc ctg cac gcc aag gcc gcg ccg gct ttg cgg 3504  
Gln Trp Glu Gly Val Ser Leu His Ala Lys Ala Ala Pro Ala Leu Arg  
1155 1160 1165

gtc cgc gtg gtc ccg gct ggt gac gat gcc aag tcc ctg ctg gtt tgt 3552  
Val Arg Val Val Pro Ala Gly Asp Asp Ala Lys Ser Leu Leu Val Cys  
1170 1175 1180

gat ggc acc ggt cga ccg gtg atc tcg gtg gac cga ctc gta ttg cgg 3600  
Asp Gly Thr Gly Arg Pro Val Ile Ser Val Asp Arg Leu Val Leu Arg  
1185 1190 1195 1200

tcg gct gcg gcc cgg cgg acc ggt gcg cgc cga cag gcc cat caa gct 3648  
Ser Ala Ala Ala Arg Arg Thr Gly Ala Arg Arg Gln Ala His Gln Ala  
1205 1210 1215

cgg ttg tac cgg ttg agc tgg cca acg gtt caa ctg ccg aca tcc gct 3696  
Arg Leu Tyr Arg Leu Ser Trp Pro Thr Val Gln Leu Pro Thr Ser Ala  
1220 1225 1230

cag cca ccg tcc ttt ttg ctt ctc ggc acc tca gaa ggc cc gct gac 3744  
 Gln Pro Pro Ser Cys Val Leu Leu Gly Thr Ser Glu Val Ser Ala Asp  
 1235 1240 1245

ata cag gtg tat ccg gac ctc ccg tcg ttg acg gct gcg ttg gat gcc 3792  
 Ile Gln Val Tyr Pro Asp Leu Arg Ser Leu Thr Ala Ala Leu Asp Ala  
 1250 1255 1260

ggt gcc gaa cca ccc ggc gtc gtc atc gca ccc acg ccc ccc ggc ggt 3840  
 Gly Ala Glu Pro Pro Gly Val Val Ile Ala Pro Thr Pro Pro Gly Gly  
 1265 1270 1275 1280

gga cga aca gcg gat gtc ccg gag acg act ccg cat gca ctc gac ctg 3888  
 Gly Arg Thr Ala Asp Val Arg Glu Thr Thr Arg His Ala Leu Asp Leu  
 1285 1290 1295

gta caa ggc tgg ctt tcc gat cag cga ctc aac gaa tcc cga ttg ctc 3936  
 Val Gln Gly Trp Leu Ser Asp Gln Arg Leu Asn Glu Ser Arg Leu Leu  
 1300 1305 1310

ctg gtg aca cag gga gca gtg gcc gtg gag ccg ggc gaa ccc gtg acc 3984  
 Leu Val Thr Gln Gly Ala Val Ala Val Glu Pro Gly Glu Pro Val Thr  
 1315 1320 1325

gat ctg gcg cag gcc gcg ctc tgg gga ctg ctg ccg tcg acg cag acc 4032  
 Asp Leu Ala Gln Ala Ala Leu Trp Gly Leu Leu Arg Ser Thr Gln Thr  
 1330 1335 1340

gaa cac cct gat cgc ttc gtc ctc gtc gat gtg cct gag ccc gcg caa 4080  
 Glu His Pro Asp Arg Phe Val Leu Val Asp Val Pro Glu Pro Ala Gln  
 1345 1350 1355 1360

ctc ctc ccc gcg ctg ccg ggg gtg ctg gcc tgc ggc gaa cct cag ctc 4128  
 Leu Leu Pro Ala Leu Pro Gly Val Leu Ala Cys Gly Glu Pro Gln Leu  
 1365 1370 1375

gcg ttg cga cgt ggc ggc gct cat gcg ccc aga ctg gct gga ctg ggc 4176  
 Ala Leu Arg Arg Gly Gly Ala His Ala Pro Arg Leu Ala Gly Leu Gly  
 1380 1385 1390

agc gat gac gtc ctg ccc gtg ccg gac ggc acc ggg tgg cga ttg gag 4224  
 Ser Asp Asp Val Leu Pro Val Pro Asp Gly Thr Gly Trp Arg Leu Glu  
 1395 1400 1405

gcc acg cgc ccg gga agc ctg gat ggg ttg gca ttg gtg gac gaa ccg 4272  
 Ala Thr Arg Pro Gly Ser Leu Asp Gly Leu Ala Leu Val Asp Glu Pro  
 1410 1415 1420

acg gcc acg gca ccg ctg ggt gac ggt gag gtc agg att gcg atg cgc 4320  
 Thr Ala Thr Ala Pro Leu Gly Asp Gly Glu Val Arg Ile Ala Met Arg  
 1425 1430 1435 1440

gcg gcc ggg gtg aac ttc cgg gat gcg ctc atc gcg ctc ggt atg tat 4368  
 Ala Ala Gly Val Asn Phe Arg Asp Ala Leu Ile Ala Leu Gly Met Tyr  
 1445 1450 1455

ccc ggt gtg gca tcg ctg ggc agt gag ggc gcc ggg gtc gtg gtg gag 4416  
 Pro Gly Val Ala Ser Leu Gly Ser Glu Gly Ala Gly Val Val Val Glu  
 1460 1465 1470

acc ggc ccc ggc gtc acc ggc ctg gca ccc ggc gac cgc gtg atg gga 4464  
 Thr Gly Pro Gly Val Thr Gly Leu Ala Pro Gly Asp Arg Val Met Gly  
 1475 1480 1485

gtg atc ccg aag gcg ttc ggg ccg ctc gcg gtc gcc gac cat cgc atg 4512  
 Met Ile Pro Lys Ala Phe Gly Pro Leu Ala Val Ala Asp His Arg Met  
 1490 1495 1500

gtg acg agg att ccc gct ggt tgg agc ttc gcg cgg gcc gca tcg gtg 4560  
 Val Thr Arg Ile Pro Ala Gly Trp Ser Phe Ala Arg Ala Ala Ser Val  
 1505 1510 1515 1520

ccg atc gtc ttt ctc acc gcc tac tac gcg ctg gtt gat ctc gcc ggg 4608  
 Pro Ile Val Phe Leu Thr Ala Tyr Tyr Ala Leu Val Asp Leu Ala Gly  
 1525 1530 1535

ttg aga cca ggg gag tcg ttg ctg gtt cat tcg gcc gcc ggt ggg gtg 4656  
 Leu Arg Pro Gly Glu Ser Leu Leu Val His Ser Ala Ala Gly Gly Val  
 1540 1545 1550

ggg atg gcc gcg atc caa ctc gcc agg cac ctc ggt gca gag gtg tac 4704  
 Gly Met Ala Ala Ile Gln Leu Ala Arg His Leu Gly Ala Glu Val Tyr  
 1555 1560 1565

gcc acc gct agc gag gac aag tgg caa gcc gtg gag ctg agc cga gaa 4752  
 Ala Thr Ala Ser Glu Asp Lys Trp Gln Ala Val Glu Leu Ser Arg Glu  
 1570 1575 1580

cac ctc gct tcg tcg cgg acg tgc gat ttc gag cag cag ttc ctc ggg 4800  
 His Leu Ala Ser Ser Arg Thr Cys Asp Phe Glu Gln Gln Phe Leu Gly  
 1585 1590 1595 1600

gca acc ggc gga cgc ggc gtc gac gtc gtg ctc aac tcc ctc gcc ggg 4848  
 Ala Thr Gly Gly Arg Gly Val Asp Val Val Leu Asn Ser Leu Ala Gly  
 1605 1610 1615

gag ttc gcc gat g	ct ctg cga atg ctg ccg cgc ggc cgt ttc	4896
Glu Phe Ala Asp Ala Ser Leu Arg Met Leu Pro Arg Gly Gly Arg Phe		
1620	1625	1630
ctg gag ttg ggg aag acg gat gtt cgt gac ccc gtc gag gtc gcc gat	4944	
Leu Glu Leu Gly Lys Thr Asp Val Arg Asp Pro Val Glu Val Ala Asp		
1635	1640	1645
gcg cat ccg ggc gtg tct tac cag gct ttc gat acc gta gag gca ggc	4992	
Ala His Pro Gly Val Ser Tyr Gln Ala Phe Asp Thr Val Glu Ala Gly		
1650	1655	1660
ccg cag cga atc ggc gag atg ctt cac gag ctg gtg gag ttg ttc gag	5040	
Pro Gln Arg Ile Gly Glu Met Leu His Glu Leu Val Glu Leu Phe Glu		
1665	1670	1675
gga cgc gtg ctg gag ccc ctg cct gtc acg gct tgg gac gtt cgg cag	5088	
Gly Arg Val Leu Glu Pro Leu Pro Val Thr Ala Trp Asp Val Arg Gln		
1685	1690	1695
gcg ccc gag gcg cta cgg cac ctg agc caa gcg cgg cat gtg gga aag	5136	
Ala Pro Glu Ala Leu Arg His Leu Ser Gln Ala Arg His Val Gly Lys		
1700	1705	1710
ctg gtg ctc acc atg cct ccg gtg tgg gac gcc gca ggc acg gtt ctg	5184	
Leu Val Leu Thr Met Pro Pro Val Trp Asp Ala Ala Gly Thr Val Leu		
1715	1720	1725
gtt acc ggc gga acg gga gca ctt ggc gca gag gtc gcc cgg cac ctc	5232	
Val Thr Gly Gly Thr Gly Ala Leu Gly Ala Glu Val Ala Arg His Leu		
1730	1735	1740
gtg atc gag cgc ggg gtg cga aac ctg gtc ctc gtc agc agg cgc ggt	5280	
Val Ile Glu Arg Gly Val Arg Asn Leu Val Leu Val Ser Arg Arg Gly		
1745	1750	1755
ccc gca gcc agt ggc gct gct gag ctc gtg gcg caa ctg acg gcc tac	5328	
Pro Ala Ala Ser Gly Ala Ala Glu Leu Val Ala Gln Leu Thr Ala Tyr		
1765	1770	1775
ggt gcc gag gtt tcc ttg cag gct tgc gat gtc gcc gat cgt gag acc	5376	
Gly Ala Glu Val Ser Leu Gln Ala Cys Asp Val Ala Asp Arg Glu Thr		
1780	1785	1790
ttg gcg aag gtg ctt gcc agc atc ccg gac gag cat ccg ttg acc gcc	5424	
Leu Ala Lys Val Leu Ala Ser Ile Pro Asp Glu His Pro Leu Thr Ala		
1795	1800	1805

gtg gtg cac gcg gct ggt gtt ctc gac gac gga gtg tcc gaa tcg ctc	5472
Val Val His Ala Ala Gly Val Leu Asp Asp Gly Val Ser Glu Ser Leu	
1810 1815 1820	
acc gtg gag cgg ctg gac cag gtt ctg cgc ccg aag gtc gat ggc gcg	5520
Thr Val Glu Arg Leu Asp Gln Val Leu Arg Pro Lys Val Asp Gly Ala	
1825 1830 1835 1840	
cgg aat ctg ctc gag ctg atc gac ccg gac gtg gcc ctc gtg ttg ttc	5568
Arg Asn Leu Leu Glu Leu Ile Asp Pro Asp Val Ala Leu Val Leu Phe	
1845 1850 1855	
tcg tcg gtg tcg ggt gtg ctc ggc agc ggt ggg cag ggt aac tac gcg	5616
Ser Ser Val Ser Gly Val Leu Gly Ser Gly Gly Gln Gly Asn Tyr Ala	
1860 1865 1870	
gcg gcc aac tcc ttc ctc gac gca ttg gcg cag caa agg cag tcg cgc	5664
Ala Ala Asn Ser Phe Leu Asp Ala Leu Ala Gln Gln Arg Gln Ser Arg	
1875 1880 1885	
ggc cta ccg acg aga tca ttg gcc tgg ggg ccc tgg gcg gaa cat ggc	5712
Gly Leu Pro Thr Arg Ser Leu Ala Trp Gly Pro Trp Ala Glu His Gly	
1890 1895 1900	
atg gcc agc acc ttg cgc gaa gcc gag cag gat cga ttg gcg cga tct	5760
Met Ala Ser Thr Leu Arg Glu Ala Glu Gln Asp Arg Leu Ala Arg Ser	
1905 1910 1915 1920	
ggg ttg ctg ccg atc tcg acc gag gag ggg ttg tcc cag ttc gac gcc	5808
Gly Leu Leu Pro Ile Ser Thr Glu Glu Gly Leu Ser Gln Phe Asp Ala	
1925 1930 1935	
gcg tgc ggc ggc gcg cat acc gtg gtg gcg ccg gtt cga ttc agc cgc	5856
Ala Cys Gly Gly Ala His Thr Val Val Ala Pro Val Arg Phe Ser Arg	
1940 1945 1950	
ttg tcc gac ggg aac gcg atc aag ttc tcc gtc ctg caa ggt ttg gtc	5904
Leu Ser Asp Gly Asn Ala Ile Lys Phe Ser Val Leu Gln Gly Leu Val	
1955 1960 1965	
ggg ccg cat cgc gtc aac aaa gcg gcg act gcg gat gat gcc gag agc	5952
Gly Pro His Arg Val Asn Lys Ala Ala Thr Ala Asp Asp Ala Glu Ser	
1970 1975 1980	
ctc cgg aaa cgg ttg gga cgc ttg ccg gat gca gaa caa cat cgg att	6000
Leu Arg Lys Arg Leu Gly Arg Leu Pro Asp Ala Glu Gln His Arg Ile	
1985 1990 1995 2000	

ctg ctg gac ctc ggc atg cat gtg gcg gca gtg ccca gga ttc gcc 6048  
 Leu Leu Asp Leu Val Arg Met His Val Ala Ala Val Leu Gly Phe Ala  
 2005 2010 2015

ggt tct cag gag atc acc gcg gac ggc acg ttc aag gtg ctg ggc ttc 6096  
 Gly Ser Gln Glu Ile Thr Ala Asp Gly Thr Phe Lys Val Leu Gly Phe  
 2020 2025 2030

gac tcg ttg acc gtg gtc gag ttg cgc aac cgg atc aac ggg gcg acg 6144  
 Asp Ser Leu Thr Val Val Glu Leu Arg Asn Arg Ile Asn Gly Ala Thr  
 2035 2040 2045

ggg ctg cga ctg ccc gcc acc ctg gtg ttc aac tac ccg acg ccg gat 6192  
 Gly Leu Arg Leu Pro Ala Thr Leu Val Phe Asn Tyr Pro Thr Pro Asp  
 2050 2055 2060

gcg ctc gcc gcg cac ctc gtc acc gcg ctg tcc gca gac cgc ctg gcc 6240  
 Ala Leu Ala Ala His Leu Val Thr Ala Leu Ser Ala Asp Arg Leu Ala  
 2065 2070 2075 2080

ggg aca ttc gag gaa ctc gac agg tgg gcg gcg aac ctg ccc acg ctg 6288  
 Gly Thr Phe Glu Glu Leu Asp Arg Trp Ala Ala Asn Leu Pro Thr Leu  
 2085 2090 2095

gcc agg gat gag gcc acg cgg gcg cag atc acc acc cgg cta cag gcg 6336  
 Ala Arg Asp Glu Ala Thr Arg Ala Gln Ile Thr Thr Arg Leu Gln Ala  
 2100 2105 2110

atc ttg cag agc ctg gcg gac gtg tcc ggc gga acc ggc ggc ggc tcc 6384  
 Ile Leu Gln Ser Leu Ala Asp Val Ser Gly Gly Thr Gly Gly Gly Ser  
 2115 2120 2125

gtg ccg gac cgg ctc aga tcg gcc acg gac gac gag ctt ttc caa ctc 6432  
 Val Pro Asp Arg Leu Arg Ser Ala Thr Asp Asp Glu Leu Phe Gln Leu  
 2130 2135 2140

ctc gac aac gat ctc gaa ctt ccc tga 6459  
 Leu Asp Asn Asp Leu Glu Leu Pro  
 2145 2150

<210> 44  
 <211> 2152  
 <212> PRT  
 <213> Saccharopolyspora spinosa

<400> 44  
 Met Thr Val Thr Thr Ser Tyr Glu Glu Val Val Glu Ala Leu Arg Ala

1

5

10

15

Ser Leu Lys Glu Asn Glu Arg Leu Arg Arg Gly Arg Asp Arg Phe Ser  
20 25 30

Ala Glu Lys Asp Asp Pro Ile Ala Ile Val Ala Met Ser Cys Arg Tyr  
35 40 45

Pro Gly Gln Val Ser Ser Pro Glu Asp Leu Trp Gln Leu Ala Ala Gly  
50 55 60

Gly Val Asp Ala Ile Ser Glu Val Pro Gly Asp Arg Gly Trp Asp Leu  
65 70 75 80

Asp Gly Val Phe Val Pro Asp Ser Asp Arg Pro Gly Thr Ser Tyr Ala  
85 90 95

Gly Ala Gly Gly Phe Leu Gln Gly Val Ser Glu Phe Asp Ala Gly Phe  
100 105 110

Phe Gly Ile Ser Pro Arg Glu Ala Leu Ala Met Asp Pro Gln Gln Arg  
115 120 125

Leu Leu Leu Glu Val Ala Trp Glu Val Phe Glu Arg Ala Gly Leu Glu  
130 135 140

Gln Arg Ser Thr Arg Gly Ser Arg Val Gly Val Phe Val Gly Thr Asn  
145 150 155 160

Gly Gln Asp Tyr Ala Ser Trp Leu Arg Thr Pro Pro Pro Ala Val Ala  
165 170 175

Gly His Val Leu Thr Gly Gly Ala Ala Ala Val Leu Ser Gly Arg Val  
180 185 190

Ala Tyr Ser Phe Gly Phe Glu Gly Pro Ala Val Thr Val Asp Thr Ala  
195 200 205

Cys Ser Ser Ser Leu Val Ala Leu His Leu Ala Gly Gln Ala Leu Arg  
210 215 220

Ala Gly Glu Cys Asp Leu Ala Leu Ala Gly Gly Val Thr Val Met Ser  
225 230 235 240

Thr Pro Lys Val Phe Leu Glu Phe Ser Arg Gln Arg Gly Leu Ala Pro  
245 250 255

Asp Gly Arg Cys Lys Ser Phe Ala Ala Gly Ala Asp Gly Thr Gly Trp

260

265

270

Gly Glu Gly Ala Gly Leu Leu Leu Leu Glu Arg Leu Ser Asp Ala Arg  
 275 280 285

Arg Asn Gly His Glu Val Leu Ala Val Val Arg Gly Ser Ala Val Asn  
 290 295 300

Gln Asp Gly Ala Ser Asn Gly Leu Thr Ala Pro Asn Gly Ser Ser Gln  
 305 310 315 320

Gln Arg Val Ile Thr Gln Ala Leu Ala Ser Ala Gly Leu Ser Val Ser  
 325 330 335

Asp Val Asp Ala Val Glu Ala His Gly Thr Gly Thr Arg Leu Gly Asp  
 340 345 350

Pro Ile Glu Ala Gln Ala Leu Ile Ala Thr Tyr Gly Arg Asp Arg Asp  
 355 360 365

Pro Gly Arg Pro Leu Trp Leu Gly Ser Val Lys Ser Asn Ile Gly His  
 370 375 380

Thr Gln Ala Ala Ala Gly Val Ala Gly Val Ile Lys Met Val Met Ala  
 385 390 395 400

Met Arg His Gly Gln Leu Pro Arg Thr Leu His Val Glu Ser Pro Ser  
 405 410 415

Pro Glu Val Asp Trp Ser Ala Gly Thr Val Gln Leu Leu Thr Glu Asn  
 420 425 430

Thr Pro Trp Pro Arg Ser Gly Arg Val Arg Arg Val Gly Val Ser Ser  
 435 440 445

Phe Gly Ile Ser Gly Thr Asn Ala His Val Ile Leu Glu Gln Pro Pro  
 450 455 460

Gly Val Pro Ser Gln Ser Ala Gly Pro Gly Ser Gly Ser Val Val Asp  
 465 470 475 480

Val Pro Val Val Pro Trp Met Val Ser Gly Lys Thr Pro Glu Ala Leu  
 485 490 495

Ser Ala Gln Ala Thr Ala Leu Met Thr Tyr Leu Asp Glu Arg Pro Asp  
 500 505 510

Val Ser Ser Leu Asp Val Gly Tyr Ser Leu Ala Leu Thr Arg Ser Ala



515

520

525

Leu Asp Glu Arg Ala Val Val Leu Gly Ser Asp Arg Glu Thr Leu Leu  
 530 535 540

Cys Gly Val Lys Ala Leu Ser Ala Gly His Glu Ala Ser Gly Leu Val  
 545 550 555 560

Thr Gly Ser Val Gly Ala Gly Gly Arg Ile Gly Phe Val Phe Ser Gly  
 565 570 575

Gln Gly Gly Gln Trp Leu Gly Met Gly Arg Gly Leu Tyr Arg Ala Phe  
 580 585 590

Pro Val Phe Ala Ala Ala Phe Asp Glu Ala Cys Ala Glu Leu Asp Ala  
 595 600 605

His Leu Gly Gln Glu Ile Gly Val Arg Glu Val Val Ser Gly Ser Asp  
 610 615 620

Ala Gln Leu Leu Asp Arg Thr Leu Trp Ala Gln Ser Gly Leu Phe Ala  
 625 630 635 640

Leu Gln Val Gly Leu Leu Lys Leu Leu Asp Ser Trp Gly Val Arg Pro  
 645 650 655

Ser Val Val Leu Gly His Ser Val Gly Glu Leu Ala Ala Ala Phe Ala  
 660 665 670

Ala Gly Val Val Ser Leu Ser Gly Ala Ala Arg Leu Val Ala Gly Arg  
 675 680 685

Ala Arg Leu Met Gln Ala Leu Pro Ser Gly Gly Gly Met Leu Ala Val  
 690 695 700

Pro Ala Gly Glu Glu Leu Leu Trp Ser Leu Leu Ala Asp Gln Gly Asp  
 705 710 715 720

Arg Val Gly Ile Ala Ala Val Asn Ala Ala Gly Ser Val Val Leu Ser  
 725 730 735

Gly Asp Arg Asp Val Leu Asp Asp Leu Ala Gly Arg Leu Asp Gly Gln  
 740 745 750

Gly Ile Arg Ser Arg Trp Leu Arg Val Ser His Ala Phe His Ser Tyr  
 755 760 765

Arg Met Asp Pro Met Leu Ala Glu Phe Ala Glu Leu Ala Arg Thr Val

770

775

780

Asp Tyr Arg Arg Cys Glu Val Pro Ile Val Ser Thr Leu Thr Gly Asp  
 785 790 795 800

Leu Asp Asp Ala Gly Arg Met Ser Gly Pro Asp Tyr Trp Val Arg Gln  
 805 810 815

Val Arg Glu Pro Val Arg Phe Ala Asp Gly Val Gln Ala Leu Val Glu  
 820 825 830

His Asp Val Ala Thr Val Val Glu Leu Gly Pro Asp Gly Ala Leu Ser  
 835 840 845

Ala Leu Ile Gln Glu Cys Val Ala Ala Ser Asp His Ala Gly Arg Leu  
 850 855 860

er Ala Val Pro Ala Met Arg Arg Asn Gln Asp Glu Ala Gln Lys Val  
 865 870 875 880

Met Thr Ala Leu Ala His Val His Val Arg Gly Gly Ala Val Asp Trp  
 885 890 895

Arg Ser Phe Phe Ala Gly Thr Gly Ala Lys Gln Ile Glu Leu Pro Thr  
 900 905 910

Tyr Ala Phe Gln Arg Gln Arg Tyr Trp Leu Val Pro Ser Asp Ser Gly  
 915 920 925

Asp Val Thr Gly Ala Gly Leu Ala Gly Ala Glu His Pro Leu Leu Gly  
 930 935 940

Ala Val Val Pro Val Ala Gly Gly Asp Glu Val Leu Leu Thr Gly Arg  
 945 950 955 960

Ile Ser Val Arg Thr His Pro Trp Leu Ala Glu His Arg Val Leu Gly  
 965 970 975

Glu Val Ile Val Ala Gly Thr Ala Leu Leu Glu Ile Ala Leu His Ala  
 980 985 990

Gly Glu Arg Leu Gly Cys Glu Arg Val Glu Glu Leu Thr Leu Glu Ala  
 995 1000 1005

Pro Leu Val Leu Pro Glu Arg Gly Ala Ile Gln Val Gln Leu Arg Val  
 1010 1015 1020

Gly Ala Pro Glu Asn Ser Gly Arg Arg Pro Met Ala Leu Tyr Ser Arg

025                      1030                      1035                      1040  
 Pro Glu Gly Ala Ala Glu His Asp Trp Thr Arg His Ala Thr Gly Arg  
                          1045                                      1050                                      1055  
 Leu Ala Pro Gly Arg Gly Glu Ala Ala Gly Asp Leu Ala Asp Trp Pro  
                          1060                                      1065                                      1070  
 Ala Pro Gly Ala Leu Pro Val Asp Leu Asp Glu Phe Tyr Arg Asp Leu  
                          1075                                      1080                                      1085  
 Ala Glu Leu Gly Leu Glu Tyr Gly Pro Ile Phe Gln Gly Leu Lys Ala  
                          1090                                      1095                                      1100  
 Ala Trp Arg Gln Gly Asp Glu Val Tyr Ala Glu Ala Ala Leu Pro Gly  
                          1105                                      1110                                      1115                                      1120  
 Thr Glu Asp Ser Gly Phe Gly Val His Pro Ala Leu Leu Asp Ala Ala  
                          1125                                      1130                                      1135  
 Leu His Ala Thr Ala Val Arg Asp Met Asp Asp Ala Arg Leu Pro Phe  
                          1140                                      1145                                      1150  
 Gln Trp Glu Gly Val Ser Leu His Ala Lys Ala Ala Pro Ala Leu Arg  
                          1155                                      1160                                      1165  
 Val Arg Val Val Pro Ala Gly Asp Asp Ala Lys Ser Leu Leu Val Cys  
                          1170                                      1175                                      1180  
 Asp Gly Thr Gly Arg Pro Val Ile Ser Val Asp Arg Leu Val Leu Arg  
                          1185                                      1190                                      1195                                      1200  
 Ser Ala Ala Ala Arg Arg Thr Gly Ala Arg Arg Gln Ala His Gln Ala  
                          1205                                      1210                                      1215  
 Arg Leu Tyr Arg Leu Ser Trp Pro Thr Val Gln Leu Pro Thr Ser Ala  
                          1220                                      1225                                      1230  
 Gln Pro Pro Ser Cys Val Leu Leu Gly Thr Ser Glu Val Ser Ala Asp  
                          1235                                      1240                                      1245  
 Ile Gln Val Tyr Pro Asp Leu Arg Ser Leu Thr Ala Ala Leu Asp Ala  
                          1250                                      1255                                      1260  
 Gly Ala Glu Pro Pro Gly Val Val Ile Ala Pro Thr Pro Pro Gly Gly  
                          1265                                      1270                                      1275                                      1280  
 Gly Arg Thr Ala Asp Val Arg Glu Thr Thr Arg His Ala Leu Asp Leu

12

1290

1295

Val Gln Gly Trp Leu Ser Asp Gln Arg Leu Asn Glu Ser Arg Leu Leu  
1300 1305 1310

Leu Val Thr Gln Gly Ala Val Ala Val Glu Pro Gly Glu Pro Val Thr  
1315 1320 1325

Asp Leu Ala Gln Ala Ala Leu Trp Gly Leu Leu Arg Ser Thr Gln Thr  
1330 1335 1340

Glu His Pro Asp Arg Phe Val Leu Val Asp Val Pro Glu Pro Ala Gln  
345 1350 1355 1360

Leu Leu Pro Ala Leu Pro Gly Val Leu Ala Cys Gly Glu Pro Gln Leu  
1365 1370 1375

la Leu Arg Arg Gly Gly Ala His Ala Pro Arg Leu Ala Gly Leu Gly  
1380 1385 1390

Ser Asp Asp Val Leu Pro Val Pro Asp Gly Thr Gly Trp Arg Leu Glu  
1395 1400 1405

Ala Thr Arg Pro Gly Ser Leu Asp Gly Leu Ala Leu Val Asp Glu Pro  
1410 1415 1420

Thr Ala Thr Ala Pro Leu Gly Asp Gly Glu Val Arg Ile Ala Met Arg  
425 1430 1435 1440

Ala Ala Gly Val Asn Phe Arg Asp Ala Leu Ile Ala Leu Gly Met Tyr  
1445 1450 1455

Pro Gly Val Ala Ser Leu Gly Ser Glu Gly Ala Gly Val Val Val Glu  
1460 1465 1470

Thr Gly Pro Gly Val Thr Gly Leu Ala Pro Gly Asp Arg Val Met Gly  
1475 1480 1485

Met Ile Pro Lys Ala Phe Gly Pro Leu Ala Val Ala Asp His Arg Met  
1490 1495 1500

Val Thr Arg Ile Pro Ala Gly Trp Ser Phe Ala Arg Ala Ala Ser Val  
505 1510 1515 1520

Pro Ile Val Phe Leu Thr Ala Tyr Tyr Ala Leu Val Asp Leu Ala Gly  
1525 1530 1535

Leu Arg Pro Gly Glu Ser Leu Leu Val His Ser Ala Ala Gly Gly Val

1540	1545	1550
Gly Met Ala Ala Ile Gln Leu Ala Arg His Leu Gly Ala Glu Val Tyr		
1555	1560	1565
Ala Thr Ala Ser Glu Asp Lys Trp Gln Ala Val Glu Leu Ser Arg Glu		
1570	1575	1580
His Leu Ala Ser Ser Arg Thr Cys Asp Phe Glu Gln Gln Phe Leu Gly		
585	1590	1595 1600
Ala Thr Gly Gly Arg Gly Val Asp Val Val Leu Asn Ser Leu Ala Gly		
1605	1610	1615
Glu Phe Ala Asp Ala Ser Leu Arg Met Leu Pro Arg Gly Gly Arg Phe		
1620	1625	1630
Leu Glu Leu Gly Lys Thr Asp Val Arg Asp Pro Val Glu Val Ala Asp		
1635	1640	1645
Ala His Pro Gly Val Ser Tyr Gln Ala Phe Asp Thr Val Glu Ala Gly		
1650	1655	1660
Pro Gln Arg Ile Gly Glu Met Leu His Glu Leu Val Glu Leu Phe Glu		
665	1670	1675 1680
Gly Arg Val Leu Glu Pro Leu Pro Val Thr Ala Trp Asp Val Arg Gln		
1685	1690	1695
Ala Pro Glu Ala Leu Arg His Leu Ser Gln Ala Arg His Val Gly Lys		
1700	1705	1710
Leu Val Leu Thr Met Pro Pro Val Trp Asp Ala Ala Gly Thr Val Leu		
1715	1720	1725
Val Thr Gly Gly Thr Gly Ala Leu Gly Ala Glu Val Ala Arg His Leu		
1730	1735	1740
Val Ile Glu Arg Gly Val Arg Asn Leu Val Leu Val Ser Arg Arg Gly		
745	1750	1755 1760
Pro Ala Ala Ser Gly Ala Ala Glu Leu Val Ala Gln Leu Thr Ala Tyr		
1765	1770	1775
Gly Ala Glu Val Ser Leu Gln Ala Cys Asp Val Ala Asp Arg Glu Thr		
1780	1785	1790
Leu Ala Lys Val Leu Ala Ser Ile Pro Asp Glu His Pro Leu Thr Ala		

1795

1800

1805

Val Val His Ala Ala Gly Val Leu Asp Asp Gly Val Ser Glu Ser Leu  
 1810 1815 1820

Thr Val Glu Arg Leu Asp Gln Val Leu Arg Pro Lys Val Asp Gly Ala  
 825 1830 1835 1840

Arg Asn Leu Leu Glu Leu Ile Asp Pro Asp Val Ala Leu Val Leu Phe  
 1845 1850 1855

Ser Ser Val Ser Gly Val Leu Gly Ser Gly Gly Gln Gly Asn Tyr Ala  
 1860 1865 1870

Ala Ala Asn Ser Phe Leu Asp Ala Leu Ala Gln Gln Arg Gln Ser Arg  
 1875 1880 1885

Gly Leu Pro Thr Arg Ser Leu Ala Trp Gly Pro Trp Ala Glu His Gly  
 1890 1895 1900

Met Ala Ser Thr Leu Arg Glu Ala Glu Gln Asp Arg Leu Ala Arg Ser  
 905 1910 1915 1920

Gly Leu Leu Pro Ile Ser Thr Glu Glu Gly Leu Ser Gln Phe Asp Ala  
 1925 1930 1935

Ala Cys Gly Gly Ala His Thr Val Val Ala Pro Val Arg Phe Ser Arg  
 1940 1945 1950

Leu Ser Asp Gly Asn Ala Ile Lys Phe Ser Val Leu Gln Gly Leu Val  
 1955 1960 1965

Gly Pro His Arg Val Asn Lys Ala Ala Thr Ala Asp Asp Ala Glu Ser  
 1970 1975 1980

Leu Arg Lys Arg Leu Gly Arg Leu Pro Asp Ala Glu Gln His Arg Ile  
 985 1990 1995 2000

Leu Leu Asp Leu Val Arg Met His Val Ala Ala Val Leu Gly Phe Ala  
 2005 2010 2015

Gly Ser Gln Glu Ile Thr Ala Asp Gly Thr Phe Lys Val Leu Gly Phe  
 2020 2025 2030

Asp Ser Leu Thr Val Val Glu Leu Arg Asn Arg Ile Asn Gly Ala Thr  
 2035 2040 2045

Gly Leu Arg Leu Pro Ala Thr Leu Val Phe Asn Tyr Pro Thr Pro Asp

2050

2055

2060

Ala Leu Ala Ala His Leu Val Thr Ala Leu Ser Ala Asp Arg Leu Ala  
065 2070 2075 2080

Gly Thr Phe Glu Glu Leu Asp Arg Trp Ala Ala Asn Leu Pro Thr Leu  
2085 2090 2095

Ala Arg Asp Glu Ala Thr Arg Ala Gln Ile Thr Thr Arg Leu Gln Ala  
2100 2105 2110

Ile Leu Gln Ser Leu Ala Asp Val Ser Gly Gly Thr Gly Gly Gly Ser  
2115 2120 2125

Val Pro Asp Arg Leu Arg Ser Ala Thr Asp Asp Glu Leu Phe Gln Leu  
2130 2135 2140

Leu Asp Asn Asp Leu Glu Leu Pro  
145 2150

&lt;210&gt; 45

&lt;211&gt; 9513

&lt;212&gt; DNA

&lt;213&gt; Saccharopolyspora spinosa

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(9510)

&lt;223&gt; ORF20; Polyketidsynthase

&lt;400&gt; 45

atg tcg aat gaa gag aag ctc cgg gag tac ttg cgg cgt gcg ctc gtg 48  
Met Ser Asn Glu Glu Lys Leu Arg Glu Tyr Leu Arg Arg Ala Leu Val  
1 5 10 15

gat ctg cac cag gcg cgc gag cgg ctg cac gag gcg gag tcg gga gag 96  
Asp Leu His Gln Ala Arg Glu Arg Leu His Glu Ala Glu Ser Gly Glu  
20 25 30

cgg gaa ccc atc gcg atc gtg gcg atg ggc tgc cgg tac ccg ggt ggg 144  
Arg Glu Pro Ile Ala Ile Val Ala Met Gly Cys Arg Tyr Pro Gly Gly  
35 40 45

gtg cag gac ccg gaa ggg ctg tgg aaa ctg gtc gcc tcc ggt ggc gac 192  
Val Gln Asp Pro Glu Gly Leu Trp Lys Leu Val Ala Ser Gly Gly Asp  
50 55 60

gcc atc ggt gaa ttc ccc gct gat cgt ggt tgg cac ctg gac gag ctc 240  
 Ala Ile Gly Glu Phe Pro Ala Asp Arg Gly Trp His Leu Asp Glu Leu  
 65 70 75 80

tac gat ccc gac ccg gat cag ccc gga acc tgc tac acc ccg cac ggc 288  
 Tyr Asp Pro Asp Pro Asp Gln Pro Gly Thr Cys Tyr Thr Arg His Gly  
 85 90 95

ggc ttc ctg cac gac gcc ggc gag ttc gac gcg gga ttc ttc gac atc 336  
 Gly Phe Leu His Asp Ala Gly Glu Phe Asp Ala Gly Phe Phe Asp Ile  
 100 105 110

agc ccc cgt gag gcg ctg gcg atg gac ccg cag cag ccg ctg ctg ctg 384  
 Ser Pro Arg Glu Ala Leu Ala Met Asp Pro Gln Gln Arg Leu Leu Leu  
 115 120 125

gaa atc tcc tgg gag acc gtc gaa tcc gct ggg atg gac ccg agg tcc 432  
 Glu Ile Ser Trp Glu Thr Val Glu Ser Ala Gly Met Asp Pro Arg Ser  
 130 135 140

ttg ccg ggg agc cgc acc ggg gtg ttc gcg gga ttg atg tac gag ggc 480  
 Leu Arg Gly Ser Arg Thr Gly Val Phe Ala Gly Leu Met Tyr Glu Gly  
 145 150 155 160

tat gac acc ggc gcc cac ccg gca gga gaa ggt gtc gaa ggc tat ctg 528  
 Tyr Asp Thr Gly Ala His Arg Ala Gly Glu Gly Val Glu Gly Tyr Leu  
 165 170 175

gga acc ggc aat gcg gga agc gtc gcc tct ggt ccg gtt gcg tat gcg 576  
 Gly Thr Gly Asn Ala Gly Ser Val Ala Ser Gly Arg Val Ala Tyr Ala  
 180 185 190

ttc ggg ttc gag ggc cca gcg gtg acg gta gac acg gcg tgc tcg tcg 624  
 Phe Gly Phe Glu Gly Pro Ala Val Thr Val Asp Thr Ala Cys Ser Ser  
 195 200 205

tcg ttg gtg gcg ctg cat ttg gcg tgt cag tcg ttg ccg cag ggc gag 672  
 Ser Leu Val Ala Leu His Leu Ala Cys Gln Ser Leu Arg Gln Gly Glu  
 210 215 220

tgt gat ctg gcg ctg gcc ggt gga gtg acg gtg atg tcg acg ccg gag 720  
 Cys Asp Leu Ala Leu Ala Gly Gly Val Thr Val Met Ser Thr Pro Glu  
 225 230 235 240

agg ttc gtg gag ttc tcc cgt cag cgt ggt ctg gca ccg gat ggg ccg 768  
 Arg Phe Val Glu Phe Ser Arg Gln Arg Gly Leu Ala Pro Asp Gly Arg  
 245 250 255



tgt aag tcg ttc gcg gcg gct gcg gat gga acc ggt tgg ggt gag ggt	816
Cys Lys Ser Phe Ala Ala Ala Ala Asp Gly Thr Gly Trp Gly Glu Gly	
260 265 270	

gcc ggt ttg gtg ttg ctg gag cgg ctg tca gac gcc agg cgg aac ggg	864
Ala Gly Leu Val Leu Leu Glu Arg Leu Ser Asp Ala Arg Arg Asn Gly	
275 280 285	

cat cgg gta ctg gcg gtt gtt cgt ggt agc gcg gtg aat cag gac ggt	912
His Arg Val Leu Ala Val Val Arg Gly Ser Ala Val Asn Gln Asp Gly	
290 295 300	

gcg tcg aac gga ttg acg gcc ccg aac ggg ctg gcc cag gag cgg gtc	960
Ala Ser Asn Gly Leu Thr Ala Pro Asn Gly Leu Ala Gln Glu Arg Val	
305 310 315 320	

ttc cag cag gtg ctc acg agt gcg ggg ctg tcg gcg tcc gat gtg gac	1008
Ile Gln Gln Val Leu Thr Ser Ala Gly Leu Ser Ala Ser Asp Val Asp	
325 330 335	

gct gtg gag gcg cat gga acg ggt acg cgg ctt ggt gat ccg atc gag	1056
Ala Val Glu Ala His Gly Thr Gly Thr Arg Leu Gly Asp Pro Ile Glu	
340 345 350	

gcg cag gct ctg ata gcc gcc tat gga cag gat cgg gac cgg gac cgg	1104
Ala Gln Ala Leu Ile Ala Ala Tyr Gly Gln Asp Arg Asp Arg Asp Arg	
355 360 365	

ccg ctg tgg ttg ggg tcg gtc aag tcc aac atc ggt cat acg cag gcg	1152
Pro Leu Trp Leu Gly Ser Val Lys Ser Asn Ile Gly His Thr Gln Ala	
370 375 380	

gct gcg ggc gtc gct ggt gtg atc aag atg gtc atg gcg atg cgg cac	1200
Ala Ala Gly Val Ala Gly Val Ile Lys Met Val Met Ala Met Arg His	
385 390 395 400	

ggg gag ctg ccg cgc acg ttg cac gtg gac gag ccg aat tcg cac gtg	1248
Gly Glu Leu Pro Arg Thr Leu His Val Asp Glu Pro Asn Ser His Val	
405 410 415	

gac tgg tcg gct ggt gcg gtc cga ctc ctg acc gag aac atc cgc tgg	1296
Asp Trp Ser Ala Gly Ala Val Arg Leu Leu Thr Glu Asn Ile Arg Trp	
420 425 430	

cca ggg acg ggt acg cgc cgc gct gga gtg tcg tcg ttc ggg gta agc	1344
Pro Gly Thr Gly Thr Arg Arg Ala Gly Val Ser Ser Phe Gly Val Ser	
435 440 445	

ggt acc aac gca cac gtc atc ctc gaa cac gac ccg ctc gcc gtg acc	1392
Gly Thr Asn Ala His Val Ile Leu Glu His Asp Pro Leu Ala Val Thr	
450 455 460	
gag aac gag gaa gca gcg cag tcc cca gca cct ggg atc gtg ccc tgg	1440
Glu Asn Glu Glu Ala Ala Gln Ser Pro Ala Pro Gly Ile Val Pro Trp	
465 470 475 480	
gcg ttg tcc ggg cgg tcg tcg acg gcg ctg cgg gcc cag gcc gaa cgg	1488
Ala Leu Ser Gly Arg Ser Ser Thr Ala Leu Arg Ala Gln Ala Glu Arg	
485 490 495	
ctg cgc gag ctg tgc gag cag acc gat ccc gac ccc gtc gat gtc ggt	1536
Leu Arg Glu Leu Cys Glu Gln Thr Asp Pro Asp Pro Val Asp Val Gly	
500 505 510	
tca ctg gcc gcc acg cgc acg gct tgg gag cac cga gcg gtg gtg	1584
Phe Ser Leu Ala Ala Thr Arg Thr Ala Trp Glu His Arg Ala Val Val	
515 520 525	
ctt ggt cgg gac agc gct acg ttg cgc tcc ggg ctt ggc gtt gtt gcc	1632
Leu Gly Arg Asp Ser Ala Thr Leu Arg Ser Gly Leu Gly Val Val Ala	
530 535 540	
agc ggt gaa cca gcg gtc gat gtc gtt gag ggg agc gtc ctg gac ggc	1680
Ser Gly Glu Pro Ala Val Asp Val Val Glu Gly Ser Val Leu Asp Gly	
545 550 555 560	
gag gtc gtc ttc gtc ttc ccc ggt cag ggc tgg cag tgg gcc ggt atg	1728
Glu Val Val Phe Val Phe Pro Gly Gln Gly Trp Gln Trp Ala Gly Met	
565 570 575	
gca gtc gac ctg ctg gac gct tcg ccg acg ttc gcg cgc cac atg gac	1776
Ala Val Asp Leu Leu Asp Ala Ser Pro Thr Phe Ala Arg His Met Asp	
580 585 590	
gag tgc gcc acc gcg ctg cgg agg tac gtg gac tgg tcg ttg gtc gac	1824
Glu Cys Ala Thr Ala Leu Arg Arg Tyr Val Asp Trp Ser Leu Val Asp	
595 600 605	
gtg ctg cgc gga gcg gag aac tcc cca ccg ctg gac cgg gtg gac gtg	1872
Val Leu Arg Gly Ala Glu Asn Ser Pro Pro Leu Asp Arg Val Asp Val	
610 615 620	
ctc cag ccc gcg tcc ttc gcg gtg atg gtg tcg ctc gcc gag gtg tgg	1920
Leu Gln Pro Ala Ser Phe Ala Val Met Val Ser Leu Ala Glu Val Trp	
625 630 635 640	



gcg cat ccg gtg ctg acc acg gcg gtc gag gaa acc ggt gcg gag tcg	2544
Ala His Pro Val Leu Thr Thr Ala Val Glu Glu Thr Gly Ala Glu Ser	
835 840 845	
gag acc gcc gtg gcc gcg gta ggt acc ttg cga cgt gac tcg ggc ggc	2592
Glu Thr Ala Val Ala Ala Val Gly Thr Leu Arg Arg Asp Ser Gly Gly	
850 855 860	
ctc cgg agg ttg ttg cat tcg ctg gcc gag gcg tac gtg cgc ggc gcc	2640
Leu Arg Arg Leu Leu His Ser Leu Ala Glu Ala Tyr Val Arg Gly Ala	
865 870 875 880	
acc gtg gac tgg gcc gtg gcg ttc ggg ggc gcg ggc cga cgg ctg gac	2688
Thr Val Asp Trp Ala Val Ala Phe Gly Gly Ala Gly Arg Arg Leu Asp	
885 890 895	
ctg ccg acc tac ccg ttc cag cgc cag cgg tac tgg ctg gac aag gga	2736
Leu Pro Thr Tyr Pro Phe Gln Arg Gln Arg Tyr Trp Leu Asp Lys Gly	
900 905 910	
gct gcc tcc gac gag gct cgt gcg gtc tcg gac ccg gcg gcg ggc tgg	2784
Ala Ala Ser Asp Glu Ala Arg Ala Val Ser Asp Pro Ala Ala Gly Trp	
915 920 925	
ttc tgg caa gcc gtg gcg cgc caa gac ctg aaa agc gtg tcc gat gcc	2832
Phe Trp Gln Ala Val Ala Arg Gln Asp Leu Lys Ser Val Ser Asp Ala	
930 935 940	
ctc gat ctc gac gcc gac gca ccg ctg agc gca aca ctt cca gcc ctg	2880
Leu Asp Leu Asp Ala Asp Ala Pro Leu Ser Ala Thr Leu Pro Ala Leu	
945 950 955 960	
tcc gtc tgg cac cgt cag gaa cga gaa agg gtc ttg gca gac ggt tgg	2928
Ser Val Trp His Arg Gln Glu Arg Glu Arg Val Leu Ala Asp Gly Trp	
965 970 975	
cgg tac cga gtc gac tgg gta cgg gtg gcc ccg cag ccg gtc cgg aga	2976
Arg Tyr Arg Val Asp Trp Val Arg Val Ala Pro Gln Pro Val Arg Arg	
980 985 990	
acg cgg gaa acc tgg ctc ctg gtc gtt ccc ccg ggc ggc atc gag gaa	3024
Thr Arg Glu Thr Trp Leu Leu Val Val Pro Pro Gly Gly Ile Glu Glu	
995 1000 1005	
gcg ctg gtc gaa cgg ctg acg gat gcg ttg aac acg cga ggg atc agc	3072
Ala Leu Val Glu Arg Leu Thr Asp Ala Leu Asn Thr Arg Gly Ile Ser	
1010 1015 1020	

acc ctg cgc ctc gac gtg cca ccg gcg gcg acc agt ggc gaa ctc gca 3120  
 Thr Leu Arg Leu Asp Val Pro Pro Ala Ala Thr Ser Gly Glu Leu Ala  
 1025 1030 1035 1040

acc gaa ctc cgc gcc gca gcc gac ggt gac ccg gtg aag gca atc ctg 3168  
 Thr Glu Leu Arg Ala Ala Asp Gly Asp Pro Val Lys Ala Ile Leu  
 1045 1050 1055

tcg ctc acc gcg ttg gac gag cga ccc cac ccc gaa tgc aag gac gtc 3216  
 Ser Leu Thr Ala Leu Asp Glu Arg Pro His Pro Glu Cys Lys Asp Val  
 1060 1065 1070

ccg agc ggg att gcc ttg ctg ctg aac ctg gtc aag gcg ctc ggt gaa 3264  
 Pro Ser Gly Ile Ala Leu Leu Leu Asn Leu Val Lys Ala Leu Gly Glu  
 1075 1080 1085

acc gac ctc aga att cct ctg tgg acc atc acg cgt ggt gcg gtc aag 3312  
 Ala Asp Leu Arg Ile Pro Leu Trp Thr Ile Thr Arg Gly Ala Val Lys  
 1090 1095 1100

gca ggc ccc gca gat cgg ctg ctg cgc ccg atg cag gcg caa gca tgg 3360  
 Ala Gly Pro Ala Asp Arg Leu Leu Arg Pro Met Gln Ala Gln Ala Trp  
 1105 1110 1115 1120

ggt ctg ggg cga gta gcc gca ctc gaa cac ccc gag cgc tgg ggt ggg 3408  
 Gly Leu Gly Arg Val Ala Ala Leu Glu His Pro Glu Arg Trp Gly Gly  
 1125 1130 1135

ctg atc gac ctg ccg gat tcg ctg gac ggc gac gtc ctc acg agg ctg 3456  
 Leu Ile Asp Leu Pro Asp Ser Leu Asp Gly Asp Val Leu Thr Arg Leu  
 1140 1145 1150

ggc gaa gcg ctc acc aac ggc ttg gcg gaa gac caa ctg gcg att cgc 3504  
 Gly Glu Ala Leu Thr Asn Gly Leu Ala Glu Asp Gln Leu Ala Ile Arg  
 1155 1160 1165

cag tcg ggc gtg ctg gcc ccg cga ctg gta ccc gcc ccg gcg aat cag 3552  
 Gln Ser Gly Val Leu Ala Arg Arg Leu Val Pro Ala Pro Ala Asn Gln  
 1170 1175 1180

ccc gct gga cgt aag tgg cgc ccc cga ggg agc gcg ctg atc acg ggc 3600  
 Pro Ala Gly Arg Lys Trp Arg Pro Arg Gly Ser Ala Leu Ile Thr Gly  
 1185 1190 1195 1200

gga ctc ggc gcg gtg ggc gca cag gtg gcg agg tgg ttg gcc gaa atc 3648  
 Gly Leu Gly Ala Val Gly Ala Gln Val Ala Arg Trp Leu Ala Glu Ile  
 1205 1210 1215

gga gcc gag cga atc gtg ctc acc agt cga cgg ggc aac caa gca gca Gly Ala Glu Arg Ile Val Leu Thr Ser Arg Arg Gly Asn Gln Ala Ala 1220 1225 1230	3696
ggc gcc gcc gag ctg gaa gcc gaa ctc cgg gcc ctt gga gcg caa gtg Gly Ala Ala Glu Leu Glu Ala Glu Leu Arg Ala Leu Gly Ala Gln Val 1235 1240 1245	3744
tcc atc gtg gct tgc gac gtg acc gat cgt gcc gag atg tcc gca cta Ser Ile Val Ala Cys Asp Val Thr Asp Arg Ala Glu Met Ser Ala Leu 1250 1255 1260	3792
ctg gcc gag ttc gac gtc acc gcg gtg ttc cac gcg gcc gga gtc ggt Leu Ala Glu Phe Asp Val Thr Ala Val Phe His Ala Ala Gly Val Gly 1265 1270 1275 1280	3840
ggc ctg ctg ccg ttg gcg gag acc gac cag aac ggc ctg gcc gaa ata Arg Leu Leu Pro Leu Ala Glu Thr Asp Gln Asn Gly Leu Ala Glu Ile 1285 1290 1295	3888
tgc gcg gcg aag gtc cgc ggc gct cag gtg ctg gac gaa ctg tgc gac Cys Ala Ala Lys Val Arg Gly Ala Gln Val Leu Asp Glu Leu Cys Asp 1300 1305 1310	3936
agc acc gat ctc gat gcc ttc gtc ctg ttc tcc tcg ggt gcc ggg gta Ser Thr Asp Leu Asp Ala Phe Val Leu Phe Ser Ser Gly Ala Gly Val 1315 1320 1325	3984
tgg ggc ggg ggc ggt cag ggc gct tac ggc gcg gcg aac gca ttc ttg Trp Gly Gly Gly Gly Gln Gly Ala Tyr Gly Ala Ala Asn Ala Phe Leu 1330 1335 1340	4032
gac aca ctc gcc gaa caa cgc cga gca cgc ggt ctg ccg gca acc tcg Asp Thr Leu Ala Glu Gln Arg Arg Ala Arg Gly Leu Pro Ala Thr Ser 1345 1350 1355 1360	4080
atc tcc tgg ggc agt tgg gcc ggc ggc ggc atg gcc gac ggc gcg gcg Ile Ser Trp Gly Ser Trp Ala Gly Gly Gly Met Ala Asp Gly Ala Ala 1365 1370 1375	4128
ggc gaa cac ctg cgg cga cgc ggg ata cgt ccg atg ccg gcg gcg tcg Gly Glu His Leu Arg Arg Arg Gly Ile Arg Pro Met Pro Ala Ala Ser 1380 1385 1390	4176
gcc atc ctg gct ctg cag gaa gta ctt gac cag gat gag acg tgc gtg Ala Ile Leu Ala Leu Gln Glu Val Leu Asp Gln Asp Glu Thr Cys Val 1395 1400 1405	4224

tcg atc gct gat gtg gac tgg gac cga ttc gtt ccc acg ttc gcc gcg 4272  
 Ser Ile Ala Asp Val Asp Trp Asp Arg Phe Val Pro Thr Phe Ala Ala  
 1410 1415 1420

act cgc gcc acc cgg ttg ttc gac gaa gtg ccg gcg gcg aga aag gcg 4320  
 Thr Arg Ala Thr Arg Leu Phe Asp Glu Val Pro Ala Ala Arg Lys Ala  
 1425 1430 1435 1440

atg ccc gcg aat ggg ccg gca gaa cca ggc ggc tcg ccg ttc gcc cgc 4368  
 Met Pro Ala Asn Gly Pro Ala Glu Pro Gly Gly Ser Pro Phe Ala Arg  
 1445 1450 1455

aat ctc gcg gag ctg ccg gaa gcc caa cga cgc cac gaa ctg gtg gat 4416  
 Asn Leu Ala Glu Leu Pro Glu Ala Gln Arg Arg His Glu Leu Val Asp  
 1460 1465 1470

ctg gtg tgc gcc cag gtg gca acc gtg ctc ggg cac ggc agt cgc gag 4464  
 Leu Val Cys Ala Gln Val Ala Thr Val Leu Gly His Gly Ser Arg Glu  
 1475 1480 1485

gaa gtc cag ccc gag cgg gcg ttc cgc gcg ctc ggg ttc gac tcc ctc 4512  
 Glu Val Gln Pro Glu Arg Ala Phe Arg Ala Leu Gly Phe Asp Ser Leu  
 1490 1495 1500

atg gcg gtg gat ctg cgc aat cgt ttg acc acc gcc acc ggg ttg cgc 4560  
 Met Ala Val Asp Leu Arg Asn Arg Leu Thr Thr Ala Thr Gly Leu Arg  
 1505 1510 1515 1520

ctg ccg acc aca acc gtc ttc gac tac ccg aat ccg gcc gcc ttg gcc 4608  
 Leu Pro Thr Thr Thr Val Phe Asp Tyr Pro Asn Pro Ala Ala Leu Ala  
 1525 1530 1535

gct cac ctg ctc gag gag ctg gtg ggt gat gtc gcg tcg gct gcg gtg 4656  
 Ala His Leu Leu Glu Glu Leu Val Gly Asp Val Ala Ser Ala Ala Val  
 1540 1545 1550

acc gct gcc agc gcg ccc gcg agt gac gaa ccg atc gcg atc gtc gcg 4704  
 Thr Ala Ala Ser Ala Pro Ala Ser Asp Glu Pro Ile Ala Ile Val Ala  
 1555 1560 1565

atg agc tgc cgg ttt ccg ggt ggc gcg cac tcg ccg gaa gac ctg tgg 4752  
 Met Ser Cys Arg Phe Pro Gly Gly Ala His Ser Pro Glu Asp Leu Trp  
 1570 1575 1580

cgg ctg gtc gcc gcc ggc acg gag gtg atc ggc gag ttc ccc tcc gac 4800  
 Arg Leu Val Ala Ala Gly Thr Glu Val Ile Gly Glu Phe Pro Ser Asp  
 1585 1590 1595 1600

cgg ggc tgg gat gcg gaa ggc ctt tac gat ccg gat gct tcc agg cct	4848
Arg Gly Trp Asp Ala Glu Gly Leu Tyr Asp Pro Asp Ala Ser Arg Pro	
1605 1610 1615	
gga acg acg tat gcg cgg atg gcg gga ttc ctc tac gac gcc ggt gag	4896
Gly Thr Thr Tyr Ala Arg Met Ala Gly Phe Leu Tyr Asp Ala Gly Glu	
1620 1625 1630	
ttc gat gcc gac ctg ttc ggc atc agc cca cgt gag gcg ttg gcg atg	4944
Phe Asp Ala Asp Leu Phe Gly Ile Ser Pro Arg Glu Ala Leu Ala Met	
1635 1640 1645	
gat ccg cag cag cgg ttg gtg ctc gaa atc gcc tgg gaa gcc ctc gaa	4992
Asp Pro Gln Gln Arg Leu Val Leu Glu Ile Ala Trp Glu Ala Leu Glu	
1650 1655 1660	
cgg gcc gga atc gat ccg ttg tcc ttg aag ggc agt ggg gtc ggc acg	5040
Arg Ala Gly Ile Asp Pro Leu Ser Leu Lys Gly Ser Gly Val Gly Thr	
1665 1670 1675 1680	
tac atc ggc gct gga agc cgt ggg tac gcg acg gat gtg cgg cag ttt	5088
Tyr Ile Gly Ala Gly Ser Arg Gly Tyr Ala Thr Asp Val Arg Gln Phe	
1685 1690 1695	
ccc gag gag gcg gag ggc tac ctg ctg acg ggt acc tcg gcc agt gtg	5136
Pro Glu Glu Ala Glu Gly Tyr Leu Leu Thr Gly Thr Ser Ala Ser Val	
1700 1705 1710	
ctg tcg ggt cgg gtc gcg tat tcg ttt ggt ttc gag ggt cct gcg gtg	5184
Leu Ser Gly Arg Val Ala Tyr Ser Phe Gly Phe Glu Gly Pro Ala Val	
1715 1720 1725	
acg gtg gat acg gct tgt tcg tcg tcg ttg gtg gcg ttg cat ctg gcg	5232
Thr Val Asp Thr Ala Cys Ser Ser Ser Leu Val Ala Leu His Leu Ala	
1730 1735 1740	
tgc cag tcg ttg cgt tcg ggc gag tgt gat ctg gcg ttg gcc ggt ggt	5280
Cys Gln Ser Leu Arg Ser Gly Glu Cys Asp Leu Ala Leu Ala Gly Gly	
1745 1750 1755 1760	
gtg acc gtg atg tcg acg ccg gag atg ttc gtg gag ttc tcc cgt cag	5328
Val Thr Val Met Ser Thr Pro Glu Met Phe Val Glu Phe Ser Arg Gln	
1765 1770 1775	
cgc ggt ttg gcg ccg gat ggg cgg tgc aag tcg ttc gcg gag agc gcg	5376
Arg Gly Leu Ala Pro Asp Gly Arg Cys Lys Ser Phe Ala Glu Ser Ala	
1780 1785 1790	



gac ggc acc ggc tgg ggc gaa ggc gcg ggc ctg ttg ttg ctg gag cgg	5424
Asp Gly Thr Gly Trp Gly Glu Gly Ala Gly Leu Leu Leu Leu Glu Arg	
1795 1800 1805	
ttg tcg gac gcc cac cgg aat ggg cat cgg gtg ttg gcg gtg gtt cgt	5472
Leu Ser Asp Ala His Arg Asn Gly His Arg Val Leu Ala Val Val Arg	
1810 1815 1820	
ggg tca gcg gtg aat cag gac ggc gcc tcg aac gga ctg gcg gcg ccg	5520
Gly Ser Ala Val Asn Gln Asp Gly Ala Ser Asn Gly Leu Ala Ala Pro	
1825 1830 1835 1840	
aac ggt ccg tcg cag cag cgg gtg atc aac cag gca ctc gcg aat gcg	5568
Asn Gly Pro Ser Gln Gln Arg Val Ile Asn Gln Ala Leu Ala Asn Ala	
1845 1850 1855	
ct ctt tcg gcg tcc gat gtg gat gcg gtg gag gca cat ggc acc ggg	5616
Ala Leu Ser Ala Ser Asp Val Asp Ala Val Glu Ala His Gly Thr Gly	
1860 1865 1870	
acc agg ctg ggt gat ccg atc gag gcg cag gca ttg atc gca acg tat	5664
Thr Arg Leu Gly Asp Pro Ile Glu Ala Gln Ala Leu Ile Ala Thr Tyr	
1875 1880 1885	
ggg cag gcc cgg gag cgg gat cgg ccc ttg tgg ctg ggg tcg gtc aag	5712
Gly Gln Ala Arg Glu Arg Asp Arg Pro Leu Trp Leu Gly Ser Val Lys	
1890 1895 1900	
tcg aac atc ggt cat acg cag gcc gcg gcg ggt gtt gcc ggt gtg atc	5760
Ser Asn Ile Gly His Thr Gln Ala Ala Ala Gly Val Ala Gly Val Ile	
1905 1910 1915 1920	
aag atg gtg atg gcc atg cgg cac ggg cag ctg ccc gcc tcg ctg cac	5808
Lys Met Val Met Ala Met Arg His Gly Gln Leu Pro Ala Ser Leu His	
1925 1930 1935	
gcg gat gag ccc acg tcg gag gtc gat tgg tcg tcg ggg gcg gtc cgg	5856
Ala Asp Glu Pro Thr Ser Glu Val Asp Trp Ser Ser Gly Ala Val Arg	
1940 1945 1950	
ctc ctc gcc gaa cag gta cct tgg ccg gag tct gac cgt gtt cgt cgg	5904
Leu Leu Ala Glu Gln Val Pro Trp Pro Glu Ser Asp Arg Val Arg Arg	
1955 1960 1965	
gtg ggg gtt tcg tcg ttc ggg atc agc ggc acc aac gca cat gtg atc	5952
Val Gly Val Ser Ser Phe Gly Ile Ser Gly Thr Asn Ala His Val Ile	
1970 1975 1980	

ctc gaa caa gct acg aat gcg cca gat agt aca gcg gag acg gac aaa 6000  
 Leu Glu Gln Ala Thr Asn Ala Pro Asp Ser Thr Ala Glu Thr Asp Lys  
 1985 1990 1995 2000

aca gaa tcc gga tct act gtc gat att ccg gtc gtt ccc tgg ttg gtg 6048  
 Thr Glu Ser Gly Ser Thr Val Asp Ile Pro Val Val Pro Trp Leu Val  
 2005 2010 2015

tcg gga aag acg acg gat tcc ctg cgg gga caa gcc gaa cga gtc ttg 6096  
 Ser Gly Lys Thr Thr Asp Ser Leu Arg Gly Gln Ala Glu Arg Val Leu  
 2020 2025 2030

tct cag gtc gag tcc cgg ccg gag cag cgt tcg ctg gat gtt gcc tac 6144  
 Ser Gln Val Glu Ser Arg Pro Glu Gln Arg Ser Leu Asp Val Ala Tyr  
 2035 2040 2045

tcg ctt gct tct ggc cga gcc gcg ctg gat gaa cgc gct gtc gtg ctg 6192  
 Ser Leu Ala Ser Gly Arg Ala Ala Leu Asp Glu Arg Ala Val Val Leu  
 2050 2055 2060

ggc ggc gac cgc ggt gag ctg gtt gct gga ctg gcg gcg ttg gcc gcc 6240  
 Gly Ala Asp Arg Gly Glu Leu Val Ala Gly Leu Ala Ala Leu Ala Ala  
 2065 2070 2075 2080

ggc cag gag gct tct ggg gtg atc agc gga act cgt gct tct gct cgg 6288  
 Gly Gln Glu Ala Ser Gly Val Ile Ser Gly Thr Arg Ala Ser Ala Arg  
 2085 2090 2095

ttc ggg ttc gtg ttc tcg ggg cag ggt ggt cag tgg ttg ggg atg ggc 6336  
 Phe Gly Phe Val Phe Ser Gly Gln Gly Gly Gln Trp Leu Gly Met Gly  
 2100 2105 2110

aga gcg ctc tac tcg aag ttt ccg gtg ttc gct gct gcg ttt gat gag 6384  
 Arg Ala Leu Tyr Ser Lys Phe Pro Val Phe Ala Ala Ala Phe Asp Glu  
 2115 2120 2125

gct tgc gcc gag ttg gag gca cat ctg ggg gaa gac cgc cgg gtt cgg 6432  
 Ala Cys Ala Glu Leu Glu Ala His Leu Gly Glu Asp Arg Arg Val Arg  
 2130 2135 2140

gat gtg gtc ttc ggt tcc gat gcg cag ctg ctg gat cag acg ctg tgg 6480  
 Asp Val Val Phe Gly Ser Asp Ala Gln Leu Leu Asp Gln Thr Leu Trp  
 2145 2150 2155 2160

gcg cag tcg ggt ctg ttc gcg ctg caa gcc ggc ctc ttg ggg ctg ctg 6528  
 Ala Gln Ser Gly Leu Phe Ala Leu Gln Ala Gly Leu Leu Gly Leu Leu  
 2165 2170 2175

ggt tgc tgg ggc gtt cgg ccg gat gtg gtg atg ggg cat tgc gtc ggg Gly Ser Trp Gly Val Arg Pro Asp Val Val Met Gly His Ser Val Gly 2180 2185 2190	6576
gag ttg gcc gcc gcg ttt gcg gct ggc gtg ttg tgc ttg cgg gat gcg Glu Leu Ala Ala Ala Phe Ala Ala Gly Val Leu Ser Leu Arg Asp Ala 2195 2200 2205	6624
gct cgg ttg gtg gcc gcg cgc gcc cgg ttg atg caa gcc ctg ccc tct Ala Arg Leu Val Ala Ala Arg Ala Arg Leu Met Gln Ala Leu Pro Ser 2210 2215 2220	6672
gac ggc gcg atg ttg gcg gtg gct gct ggt gaa gac ctt gtt cgg cca Asp Gly Ala Met Leu Ala Val Ala Ala Gly Glu Asp Leu Val Arg Pro 2225 2230 2235 2240	6720
ctg ctg gcc ggt cgg gag gag tcc gtg agc gtc gcc gcg ctc aat gcc Leu Leu Ala Gly Arg Glu Glu Ser Val Ser Val Ala Ala Leu Asn Ala 2245 2250 2255	6768
ccc ggt tgc gtg gtg ttg tgc ggc gat cgg gag gtg ctg gcc agc atc Pro Gly Ser Val Val Leu Ser Gly Asp Arg Glu Val Leu Ala Ser Ile 2260 2265 2270	6816
gtc ggc cgg ctg acc gag ctc cga gtc cgg acg cgg cgc ttg cgg gtc Val Gly Arg Leu Thr Glu Leu Arg Val Arg Thr Arg Arg Leu Arg Val 2275 2280 2285	6864
tcc cat gct ttt cat tgc cac cgg atg gac ccg atg ttg ggc gag ttc Ser His Ala Phe His Ser His Arg Met Asp Pro Met Leu Gly Glu Phe 2290 2295 2300	6912
gcc cag atc gcc gag tct gcg gag ttc ggt aag cca acg aca ccg ctt Ala Gln Ile Ala Glu Ser Ala Glu Phe Gly Lys Pro Thr Thr Pro Leu 2305 2310 2315 2320	6960
gtg tgc acg ttg acg ggt gag ctc gac aga gcc gcg gaa atg agc aca Val Ser Thr Leu Thr Gly Glu Leu Asp Arg Ala Ala Glu Met Ser Thr 2325 2330 2335	7008
cca ggg tat tgg gtg cgc cag gcg cgt gaa ccc gtc cgt ttc gcc gac Pro Gly Tyr Trp Val Arg Gln Ala Arg Glu Pro Val Arg Phe Ala Asp 2340 2345 2350	7056
ggt gtc cag gcc ctg gca gcg cag ggc ata ggc acg gtc gtc gag ctc Gly Val Gln Ala Leu Ala Ala Gln Gly Ile Gly Thr Val Val Glu Leu 2355 2360 2365	7104

ggc ccg gac gga acg ctg gcg gca ctg gtt cgg gag tgt gcg acc gag	7152
Gly Pro Asp Gly Thr Leu Ala Ala Leu Val Arg Glu Cys Ala Thr Glu	
2370 2375 2380	
tcc gat cgg gtt ggg cgg att tcg tcg atc cca ctg atg cgc agg gag	7200
Ser Asp Arg Val Gly Arg Ile Ser Ser Ile Pro Leu Met Arg Arg Glu	
2385 2390 2395 2400	
cgg gac gag acc cgt tcg gtg atg aca gcc ctg gcg cat ctc cac acc	7248
Arg Asp Glu Thr Arg Ser Val Met Thr Ala Leu Ala His Leu His Thr	
2405 2410 2415	
cgt ggt ggt gag gtg gac tgg cag gcg ttt ttc gcc ggt acc ggc gct	7296
Arg Gly Gly Glu Val Asp Trp Gln Ala Phe Phe Ala Gly Thr Gly Ala	
2420 2425 2430	
ggc cag ctc gag ttg cca acg tat gcc ttc caa cga cag cac tac tgg	7344
Arg Gln Leu Glu Leu Pro Thr Tyr Ala Phe Gln Arg Gln His Tyr Trp	
2435 2440 2445	
atc gag tcc agt gcg cgg cca gca cgc gac cgc gca gac atc ggc gag	7392
Ile Glu Ser Ser Ala Arg Pro Ala Arg Asp Arg Ala Asp Ile Gly Glu	
2450 2455 2460	
gtg gcg gaa cag ttc tgg acc gcg gtt gac caa ggc gat ctg gca acg	7440
Val Ala Glu Gln Phe Trp Thr Ala Val Asp Gln Gly Asp Leu Ala Thr	
2465 2470 2475 2480	
ttg gtc gcc gct ctg gat ctt ggg gcg gac gac gac aca tgc gca tcg	7488
Leu Val Ala Ala Leu Asp Leu Gly Ala Asp Asp Asp Thr Cys Ala Ser	
2485 2490 2495	
ttg agc gat gta ttg ccg gcg ttg tcc tcc tgg cga agc gga ctc cgc	7536
Leu Ser Asp Val Leu Pro Ala Leu Ser Ser Trp Arg Ser Gly Leu Arg	
2500 2505 2510	
aac cgt tcg ctc gtc gat tcc tgc cgg tac cga atc agt tgg cat tcc	7584
Asn Arg Ser Leu Val Asp Ser Cys Arg Tyr Arg Ile Ser Trp His Ser	
2515 2520 2525	
tct cgg gag gtg ccg gcc ccg aag att tcc ggt acc tgg ctg ttg gtc	7632
Ser Arg Glu Val Pro Ala Pro Lys Ile Ser Gly Thr Trp Leu Leu Val	
2530 2535 2540	
gtg ccc ggt gct gcg gat gac gga ttg gtc acg gct ttg acg agt tca	7680
Val Pro Gly Ala Ala Asp Asp Gly Leu Val Thr Ala Leu Thr Ser Ser	
2545 2550 2555 2560	

ctg gtc gga ggc ggc gcc gag gtc gtc cgg atc ggc ctg tcc gaa gag	7728
Leu Val Gly Gly Gly Ala Glu Val Val Arg Ile Gly Leu Ser Glu Glu	
2565 2570 2575	
gac ccg cac cgc gag gac gtc gca cag cgg ctg gcc aat gcg ctg acg	7776
Asp Pro His Arg Glu Asp Val Ala Gln Arg Leu Ala Asn Ala Leu Thr	
2580 2585 2590	
gat gcc ggt caa ctc ggt ggc gtg ctt tcg ctg ttg ggg ctc gat gaa	7824
Asp Ala Gly Gln Leu Gly Gly Val Leu Ser Leu Leu Gly Leu Asp Glu	
2595 2600 2605	
tcg cct gct ccg gga ttc tcc tgc ttg cca act ggt ttc gcg ctg act	7872
Ser Pro Ala Pro Gly Phe Ser Cys Leu Pro Thr Gly Phe Ala Leu Thr	
2610 2615 2620	
tg cag ctt ctg cgg gcc ttg cgg aag gcc gac gtc gag gcg cct ttt	7920
Val Gln Leu Leu Arg Ala Leu Arg Lys Ala Asp Val Glu Ala Pro Phe	
2625 2630 2635 2640	
tgg gcg gtg acg cgc ggc ggc gtc gcg ttg gaa gat gta cgc gtg tct	7968
Trp Ala Val Thr Arg Gly Gly Val Ala Leu Glu Asp Val Arg Val Ser	
2645 2650 2655	
ccg gag cag gcc ctg gtc tgg ggg ctg ctg cgt gtc gcg gga ctg gag	8016
Pro Glu Gln Ala Leu Val Trp Gly Leu Leu Arg Val Ala Gly Leu Glu	
2660 2665 2670	
cac ccg gag ttc tgg ggt ggc ttg atc gac ctg cca tcg gac tgg gac	8064
His Pro Glu Phe Trp Gly Gly Leu Ile Asp Leu Pro Ser Asp Trp Asp	
2675 2680 2685	
gac cga ttg ggt gcc cgg ttg gcg ggt gtg ttg gcg gat ggt ggc gag	8112
Asp Arg Leu Gly Ala Arg Leu Ala Gly Val Leu Ala Asp Gly Gly Glu	
2690 2695 2700	
gat caa gtc gcc att cgc cgt ggt ggt gtg ttc gtg cgg cgg ttg gaa	8160
Asp Gln Val Ala Ile Arg Arg Gly Gly Val Phe Val Arg Arg Leu Glu	
2705 2710 2715 2720	
cgc gct ggt gcg tcg ggt gcc ggg tcg gtg tgg cgt cct cgg ggg acg	8208
Arg Ala Gly Ala Ser Gly Ala Gly Ser Val Trp Arg Pro Arg Gly Thr	
2725 2730 2735	
gtg ttg gtg acg ggt ggt acg ggc ggt ttg ggg gcg cat gtt gcc cgg	8256
Val Leu Val Thr Gly Gly Thr Gly Gly Leu Gly Ala His Val Ala Arg	
2740 2745 2750	



gca ttg gac cgc gat gag acc tcc ctc gtc gtg gcc gat gtc gat tgg	8880
Ala Leu Asp Arg Asp Glu Thr Ser Leu Val Val Ala Asp Val Asp Trp	
2945 2950 2955 2960	
gca cgg ttc gcc cca gcc ttc acc tcg gca cgt cga cgc ccg ctg ctg	8928
Ala Arg Phe Ala Pro Ala Phe Thr Ser Ala Arg Arg Arg Pro Leu Leu	
2965 2970 2975	
gac acc atc gac gag gcc cga gcc gca ttg gaa acc acc ggc gaa caa	8976
Asp Thr Ile Asp Glu Ala Arg Ala Ala Leu Glu Thr Thr Gly Glu Gln	
2980 2985 2990	
gcg ggc aca ggc aaa ccc gtt gag ctg acg caa cgc ctg gcc gga ctg	9024
Ala Gly Thr Gly Lys Pro Val Glu Leu Thr Gln Arg Leu Ala Gly Leu	
2995 3000 3005	
cg cgg aag gaa cgc gac gat gcg gta ttg gat ctg gtg cgg gcg gag	9072
Ser Arg Lys Glu Arg Asp Asp Ala Val Leu Asp Leu Val Arg Ala Glu	
3010 3015 3020	
acg gcg gct gtg ctg gga cgc gac gat gcc acg gcc ctg gcg cca tcg	9120
Thr Ala Ala Val Leu Gly Arg Asp Asp Ala Thr Ala Leu Ala Pro Ser	
3025 3030 3035 3040	
cgg ccg ttc cag gaa ctc gga ttc gac tcc ttg atg gcg gtg gag ctg	9168
Arg Pro Phe Gln Glu Leu Gly Phe Asp Ser Leu Met Ala Val Glu Leu	
3045 3050 3055	
cgc aac cgg ctg aac acc gcc acc ggg atc cag ctg ccc gcc agc acg	9216
Arg Asn Arg Leu Asn Thr Ala Thr Gly Ile Gln Leu Pro Ala Ser Thr	
3060 3065 3070	
att ttc gac tac ccc aat gcc gag tcg ctg tcg cgt cac ctc tgc gcc	9264
Ile Phe Asp Tyr Pro Asn Ala Glu Ser Leu Ser Arg His Leu Cys Ala	
3075 3080 3085	
gag ctt ttc cca acg gag act acc gtg gac tcg gcc ctt gcc gag ctc	9312
Glu Leu Phe Pro Thr Glu Thr Thr Val Asp Ser Ala Leu Ala Glu Leu	
3090 3095 3100	
gat cga atc gag cag cag ctc tcg atg ctc acc ggc gaa gcg cgg gca	9360
Asp Arg Ile Glu Gln Gln Leu Ser Met Leu Thr Gly Glu Ala Arg Ala	
3105 3110 3115 3120	
cgg gac cga atc gcg aca cga ctg cga gcc ctc cac gag aag tgg aac	9408
Arg Asp Arg Ile Ala Thr Arg Leu Arg Ala Leu His Glu Lys Trp Asn	
3125 3130 3135	

agc gca gct gaa gta ccg acc gga gcc gat gtc ctg agc acg ctc gat 9456  
Ser Ala Ala Glu Val Pro Thr Gly Ala Asp Val Leu Ser Thr Leu Asp  
3140 3145 3150

tcg gcg acg cac gac gag ata ttc gag ttc atc gac aac gag ctc gac 9504  
Ser Ala Thr His Asp Glu Ile Phe Glu Phe Ile Asp Asn Glu Leu Asp  
3155 3160 3165

ctg tcc tga 9513  
Leu Ser  
3170

<210> 46

<211> 3170

<212> PRT

<213> Saccharopolyspora spinosa

<400> 46

Met Ser Asn Glu Glu Lys Leu Arg Glu Tyr Leu Arg Arg Ala Leu Val  
1 5 10 15

Asp Leu His Gln Ala Arg Glu Arg Leu His Glu Ala Glu Ser Gly Glu  
20 25 30

Arg Glu Pro Ile Ala Ile Val Ala Met Gly Cys Arg Tyr Pro Gly Gly  
35 40 45

Val Gln Asp Pro Glu Gly Leu Trp Lys Leu Val Ala Ser Gly Gly Asp  
50 55 60

Ala Ile Gly Glu Phe Pro Ala Asp Arg Gly Trp His Leu Asp Glu Leu  
65 70 75 80

Tyr Asp Pro Asp Pro Asp Gln Pro Gly Thr Cys Tyr Thr Arg His Gly  
85 90 95

Gly Phe Leu His Asp Ala Gly Glu Phe Asp Ala Gly Phe Phe Asp Ile  
100 105 110

Ser Pro Arg Glu Ala Leu Ala Met Asp Pro Gln Gln Arg Leu Leu Leu  
115 120 125

Glu Ile Ser Trp Glu Thr Val Glu Ser Ala Gly Met Asp Pro Arg Ser  
130 135 140

Leu Arg Gly Ser Arg Thr Gly Val Phe Ala Gly Leu Met Tyr Glu Gly



145		150		155		160
Tyr Asp Thr Gly Ala His Arg Ala Gly Glu Gly Val Glu Gly Tyr Leu						
	165		170		175	
Gly Thr Gly Asn Ala Gly Ser Val Ala Ser Gly Arg Val Ala Tyr Ala						
	180		185		190	
Phe Gly Phe Glu Gly Pro Ala Val Thr Val Asp Thr Ala Cys Ser Ser						
	195		200		205	
Ser Leu Val Ala Leu His Leu Ala Cys Gln Ser Leu Arg Gln Gly Glu						
	210		215		220	
Cys Asp Leu Ala Leu Ala Gly Gly Val Thr Val Met Ser Thr Pro Glu						
	225		230		235	240
Arg Phe Val Glu Phe Ser Arg Gln Arg Gly Leu Ala Pro Asp Gly Arg						
	245		250		255	
Cys Lys Ser Phe Ala Ala Ala Ala Asp Gly Thr Gly Trp Gly Glu Gly						
	260		265		270	
Ala Gly Leu Val Leu Leu Glu Arg Leu Ser Asp Ala Arg Arg Asn Gly						
	275		280		285	
His Arg Val Leu Ala Val Val Arg Gly Ser Ala Val Asn Gln Asp Gly						
	290		295		300	
Ala Ser Asn Gly Leu Thr Ala Pro Asn Gly Leu Ala Gln Glu Arg Val						
	305		310		315	320
Ile Gln Gln Val Leu Thr Ser Ala Gly Leu Ser Ala Ser Asp Val Asp						
	325		330		335	
Ala Val Glu Ala His Gly Thr Gly Thr Arg Leu Gly Asp Pro Ile Glu						
	340		345		350	
Ala Gln Ala Leu Ile Ala Ala Tyr Gly Gln Asp Arg Asp Arg Asp Arg						
	355		360		365	
Pro Leu Trp Leu Gly Ser Val Lys Ser Asn Ile Gly His Thr Gln Ala						
	370		375		380	
Ala Ala Gly Val Ala Gly Val Ile Lys Met Val Met Ala Met Arg His						
	385		390		395	400
Gly Glu Leu Pro Arg Thr Leu His Val Asp Glu Pro Asn Ser His Val						

405

410

415

Asp Trp Ser Ala Gly Ala Val Arg Leu Leu Thr Glu Asn Ile Arg Trp  
420 425 430

Pro Gly Thr Gly Thr Arg Arg Ala Gly Val Ser Ser Phe Gly Val Ser  
435 440 445

Gly Thr Asn Ala His Val Ile Leu Glu His Asp Pro Leu Ala Val Thr  
450 455 460

Glu Asn Glu Glu Ala Ala Gln Ser Pro Ala Pro Gly Ile Val Pro Trp  
465 470 475 480

Ala Leu Ser Gly Arg Ser Ser Thr Ala Leu Arg Ala Gln Ala Glu Arg  
485 490 495

Leu Arg Glu Leu Cys Glu Gln Thr Asp Pro Asp Pro Val Asp Val Gly  
500 505 510

Phe Ser Leu Ala Ala Thr Arg Thr Ala Trp Glu His Arg Ala Val Val  
515 520 525

Leu Gly Arg Asp Ser Ala Thr Leu Arg Ser Gly Leu Gly Val Val Ala  
530 535 540

Ser Gly Glu Pro Ala Val Asp Val Val Glu Gly Ser Val Leu Asp Gly  
545 550 555 560

Glu Val Val Phe Val Phe Pro Gly Gln Gly Trp Gln Trp Ala Gly Met  
565 570 575

Ala Val Asp Leu Leu Asp Ala Ser Pro Thr Phe Ala Arg His Met Asp  
580 585 590

Glu Cys Ala Thr Ala Leu Arg Arg Tyr Val Asp Trp Ser Leu Val Asp  
595 600 605

Val Leu Arg Gly Ala Glu Asn Ser Pro Pro Leu Asp Arg Val Asp Val  
610 615 620

Leu Gln Pro Ala Ser Phe Ala Val Met Val Ser Leu Ala Glu Val Trp  
625 630 635 640

Arg Ser Tyr Gly Val Arg Pro Ala Ala Val Val Gly His Ser Gln Gly  
645 650 655

Glu Ile Ala Ala Ala Cys Ala Ala Gly Val Leu Pro Leu Glu Asp Ala

660	665	670
Ala Arg Leu Val Ala Leu Arg Ser Arg Ala Leu Lys Gly Leu Ser Gly		
675	680	685
Arg Gly Gly Met Ala Ser Leu Ala Cys Pro Ala Asp Glu Val Ala Ala		
690	695	700
Leu Phe Ala Gly Ser Gly Gly Arg Leu Glu Val Ala Ala Ile Asn Gly		
705	710	715 720
Pro Arg Ser Val Val Val Ser Gly Asp Leu Glu Ala Val Asp Glu Leu		
	725	730 735
Leu Ala Glu Cys Ala Glu Lys Asp Met Arg Ala Arg Arg Ile Pro Val		
	740	745 750
Asp Tyr Ala Ser His Ser Ala His Val Glu Val Val Arg Ser Pro Val		
	755	760 765
Leu Ala Ala Ala Ala Gly Val Arg His Arg Asp Gly Gln Val Pro Trp		
	770	775 780
Trp Ser Thr Val Ile Gly Asp Trp Val Asp Pro Ala Arg Leu Asp Gly		
785	790	795 800
Glu Tyr Trp Tyr Arg Asn Leu Arg Gln Pro Val Arg Phe Glu His Ala		
	805	810 815
Val Gln Gly Leu Val Glu Arg Gly Phe Gly Leu Phe Ile Glu Met Ser		
	820	825 830
Ala His Pro Val Leu Thr Thr Ala Val Glu Glu Thr Gly Ala Glu Ser		
	835	840 845
Glu Thr Ala Val Ala Ala Val Gly Thr Leu Arg Arg Asp Ser Gly Gly		
	850	855 860
Leu Arg Arg Leu Leu His Ser Leu Ala Glu Ala Tyr Val Arg Gly Ala		
865	870	875 880
Thr Val Asp Trp Ala Val Ala Phe Gly Gly Ala Gly Arg Arg Leu Asp		
	885	890 895
Leu Pro Thr Tyr Pro Phe Gln Arg Gln Arg Tyr Trp Leu Asp Lys Gly		
	900	905 910
Ala Ala Ser Asp Glu Ala Arg Ala Val Ser Asp Pro Ala Ala Gly Trp		

915

920

925

Phe Trp Gln Ala Val Ala Arg Gln Asp Leu Lys Ser Val Ser Asp Ala  
 930 935 940

Leu Asp Leu Asp Ala Asp Ala Pro Leu Ser Ala Thr Leu Pro Ala Leu  
 945 950 955 960

Ser Val Trp His Arg Gln Glu Arg Glu Arg Val Leu Ala Asp Gly Trp  
 965 970 975

Arg Tyr Arg Val Asp Trp Val Arg Val Ala Pro Gln Pro Val Arg Arg  
 980 985 990

Thr Arg Glu Thr Trp Leu Leu Val Val Pro Pro Gly Gly Ile Glu Glu  
 995 1000 1005

Ala Leu Val Glu Arg Leu Thr Asp Ala Leu Asn Thr Arg Gly Ile Ser  
 1010 1015 1020

Thr Leu Arg Leu Asp Val Pro Pro Ala Ala Thr Ser Gly Glu Leu Ala  
 1025 1030 1035 1040

Thr Glu Leu Arg Ala Ala Ala Asp Gly Asp Pro Val Lys Ala Ile Leu  
 1045 1050 1055

Ser Leu Thr Ala Leu Asp Glu Arg Pro His Pro Glu Cys Lys Asp Val  
 1060 1065 1070

Pro Ser Gly Ile Ala Leu Leu Leu Asn Leu Val Lys Ala Leu Gly Glu  
 1075 1080 1085

Ala Asp Leu Arg Ile Pro Leu Trp Thr Ile Thr Arg Gly Ala Val Lys  
 1090 1095 1100

Ala Gly Pro Ala Asp Arg Leu Leu Arg Pro Met Gln Ala Gln Ala Trp  
 1105 1110 1115 1120

Gly Leu Gly Arg Val Ala Ala Leu Glu His Pro Glu Arg Trp Gly Gly  
 1125 1130 1135

Leu Ile Asp Leu Pro Asp Ser Leu Asp Gly Asp Val Leu Thr Arg Leu  
 1140 1145 1150

Gly Glu Ala Leu Thr Asn Gly Leu Ala Glu Asp Gln Leu Ala Ile Arg  
 1155 1160 1165

Gln Ser Gly Val Leu Ala Arg Arg Leu Val Pro Ala Pro Ala Asn Gln

1170	1175	1180
Pro Ala Gly Arg Lys Trp Arg Pro Arg Gly Ser Ala Leu Ile Thr Gly		
185	1190	1195 1200
Gly Leu Gly Ala Val Gly Ala Gln Val Ala Arg Trp Leu Ala Glu Ile		
1205	1210	1215
Gly Ala Glu Arg Ile Val Leu Thr Ser Arg Arg Gly Asn Gln Ala Ala		
1220	1225	1230
Gly Ala Ala Glu Leu Glu Ala Glu Leu Arg Ala Leu Gly Ala Gln Val		
1235	1240	1245
Ser Ile Val Ala Cys Asp Val Thr Asp Arg Ala Glu Met Ser Ala Leu		
1250	1255	1260
Leu Ala Glu Phe Asp Val Thr Ala Val Phe His Ala Ala Gly Val Gly		
265	1270	1275 1280
Arg Leu Leu Pro Leu Ala Glu Thr Asp Gln Asn Gly Leu Ala Glu Ile		
1285	1290	1295
Cys Ala Ala Lys Val Arg Gly Ala Gln Val Leu Asp Glu Leu Cys Asp		
1300	1305	1310
Ser Thr Asp Leu Asp Ala Phe Val Leu Phe Ser Ser Gly Ala Gly Val		
1315	1320	1325
Trp Gly Gly Gly Gly Gln Gly Ala Tyr Gly Ala Ala Asn Ala Phe Leu		
1330	1335	1340
Asp Thr Leu Ala Glu Gln Arg Arg Ala Arg Gly Leu Pro Ala Thr Ser		
345	1350	1355 1360
Ile Ser Trp Gly Ser Trp Ala Gly Gly Gly Met Ala Asp Gly Ala Ala		
1365	1370	1375
Gly Glu His Leu Arg Arg Arg Gly Ile Arg Pro Met Pro Ala Ala Ser		
1380	1385	1390
Ala Ile Leu Ala Leu Gln Glu Val Leu Asp Gln Asp Glu Thr Cys Val		
1395	1400	1405
Ser Ile Ala Asp Val Asp Trp Asp Arg Phe Val Pro Thr Phe Ala Ala		
1410	1415	1420
Thr Arg Ala Thr Arg Leu Phe Asp Glu Val Pro Ala Ala Arg Lys Ala		

425	1430	1435	1440
Met Pro Ala Asn Gly Pro Ala Glu Pro Gly Gly Ser Pro Phe Ala Arg	1445	1450	1455
Asn Leu Ala Glu Leu Pro Glu Ala Gln Arg Arg His Glu Leu Val Asp	1460	1465	1470
Leu Val Cys Ala Gln Val Ala Thr Val Leu Gly His Gly Ser Arg Glu	1475	1480	1485
Glu Val Gln Pro Glu Arg Ala Phe Arg Ala Leu Gly Phe Asp Ser Leu	1490	1495	1500
Met Ala Val Asp Leu Arg Asn Arg Leu Thr Thr Ala Thr Gly Leu Arg	505	1510	1515
Leu Pro Thr Thr Thr Val Phe Asp Tyr Pro Asn Pro Ala Ala Leu Ala	1525	1530	1535
Ala His Leu Leu Glu Glu Leu Val Gly Asp Val Ala Ser Ala Ala Val	1540	1545	1550
Thr Ala Ala Ser Ala Pro Ala Ser Asp Glu Pro Ile Ala Ile Val Ala	1555	1560	1565
Met Ser Cys Arg Phe Pro Gly Gly Ala His Ser Pro Glu Asp Leu Trp	1570	1575	1580
Arg Leu Val Ala Ala Gly Thr Glu Val Ile Gly Glu Phe Pro Ser Asp	585	1590	1595
Arg Gly Trp Asp Ala Glu Gly Leu Tyr Asp Pro Asp Ala Ser Arg Pro	1605	1610	1615
Gly Thr Thr Tyr Ala Arg Met Ala Gly Phe Leu Tyr Asp Ala Gly Glu	1620	1625	1630
Phe Asp Ala Asp Leu Phe Gly Ile Ser Pro Arg Glu Ala Leu Ala Met	1635	1640	1645
Asp Pro Gln Gln Arg Leu Val Leu Glu Ile Ala Trp Glu Ala Leu Glu	1650	1655	1660
Arg Ala Gly Ile Asp Pro Leu Ser Leu Lys Gly Ser Gly Val Gly Thr	665	1670	1675
Tyr Ile Gly Ala Gly Ser Arg Gly Tyr Ala Thr Asp Val Arg Gln Phe			1680

1685	1690	1695
Pro Glu Glu Ala Glu Gly Tyr Leu Leu Thr Gly Thr Ser Ala Ser Val		
1700	1705	1710
Leu Ser Gly Arg Val Ala Tyr Ser Phe Gly Phe Glu Gly Pro Ala Val		
1715	1720	1725
Thr Val Asp Thr Ala Cys Ser Ser Ser Leu Val Ala Leu His Leu Ala		
1730	1735	1740
Cys Gln Ser Leu Arg Ser Gly Glu Cys Asp Leu Ala Leu Ala Gly Gly		
745	1750	1755 1760
Val Thr Val Met Ser Thr Pro Glu Met Phe Val Glu Phe Ser Arg Gln		
1765	1770	1775
Arg Gly Leu Ala Pro Asp Gly Arg Cys Lys Ser Phe Ala Glu Ser Ala		
1780	1785	1790
Asp Gly Thr Gly Trp Gly Glu Gly Ala Gly Leu Leu Leu Leu Glu Arg		
1795	1800	1805
Leu Ser Asp Ala His Arg Asn Gly His Arg Val Leu Ala Val Val Arg		
1810	1815	1820
Gly Ser Ala Val Asn Gln Asp Gly Ala Ser Asn Gly Leu Ala Ala Pro		
825	1830	1835 1840
Asn Gly Pro Ser Gln Gln Arg Val Ile Asn Gln Ala Leu Ala Asn Ala		
1845	1850	1855
Ala Leu Ser Ala Ser Asp Val Asp Ala Val Glu Ala His Gly Thr Gly		
1860	1865	1870
Thr Arg Leu Gly Asp Pro Ile Glu Ala Gln Ala Leu Ile Ala Thr Tyr		
1875	1880	1885
Gly Gln Ala Arg Glu Arg Asp Arg Pro Leu Trp Leu Gly Ser Val Lys		
1890	1895	1900
Ser Asn Ile Gly His Thr Gln Ala Ala Ala Gly Val Ala Gly Val Ile		
905	1910	1915 1920
Lys Met Val Met Ala Met Arg His Gly Gln Leu Pro Ala Ser Leu His		
1925	1930	1935
Ala Asp Glu Pro Thr Ser Glu Val Asp Trp Ser Ser Gly Ala Val Arg		

1940

1945

50

Leu Leu Ala Glu Gln Val Pro Trp Pro Glu Ser Asp Arg Val Arg Arg  
 1955 1960 1965

Val Gly Val Ser Ser Phe Gly Ile Ser Gly Thr Asn Ala His Val Ile  
 1970 1975 1980

Leu Glu Gln Ala Thr Asn Ala Pro Asp Ser Thr Ala Glu Thr Asp Lys  
 985 1990 1995 2000

Thr Glu Ser Gly Ser Thr Val Asp Ile Pro Val Val Pro Trp Leu Val  
 2005 2010 2015

Ser Gly Lys Thr Thr Asp Ser Leu Arg Gly Gln Ala Glu Arg Val Leu  
 2020 2025 2030

Ser Gln Val Glu Ser Arg Pro Glu Gln Arg Ser Leu Asp Val Ala Tyr  
 2035 2040 2045

Ser Leu Ala Ser Gly Arg Ala Ala Leu Asp Glu Arg Ala Val Val Leu  
 2050 2055 2060

Gly Ala Asp Arg Gly Glu Leu Val Ala Gly Leu Ala Ala Leu Ala Ala  
 065 2070 2075 2080

Gly Gln Glu Ala Ser Gly Val Ile Ser Gly Thr Arg Ala Ser Ala Arg  
 2085 2090 2095

Phe Gly Phe Val Phe Ser Gly Gln Gly Gly Gln Trp Leu Gly Met Gly  
 2100 2105 2110

Arg Ala Leu Tyr Ser Lys Phe Pro Val Phe Ala Ala Ala Phe Asp Glu  
 2115 2120 2125

Ala Cys Ala Glu Leu Glu Ala His Leu Gly Glu Asp Arg Arg Val Arg  
 2130 2135 2140

Asp Val Val Phe Gly Ser Asp Ala Gln Leu Leu Asp Gln Thr Leu Trp  
 145 2150 2155 2160

Ala Gln Ser Gly Leu Phe Ala Leu Gln Ala Gly Leu Leu Gly Leu Leu  
 2165 2170 2175

Gly Ser Trp Gly Val Arg Pro Asp Val Val Met Gly His Ser Val Gly  
 2180 2185 2190

Glu Leu Ala Ala Ala Phe Ala Ala Gly Val Leu Ser Leu Arg Asp Ala



2195

2200

2205

Ala Arg Leu Val Ala Ala Arg Ala Arg Leu Met Gln Ala Leu Pro Ser  
2210 2215 2220

Asp Gly Ala Met Leu Ala Val Ala Ala Gly Glu Asp Leu Val Arg Pro  
225 2230 2235 2240

Leu Leu Ala Gly Arg Glu Glu Ser Val Ser Val Ala Ala Leu Asn Ala  
2245 2250 2255

Pro Gly Ser Val Val Leu Ser Gly Asp Arg Glu Val Leu Ala Ser Ile  
2260 2265 2270

Val Gly Arg Leu Thr Glu Leu Arg Val Arg Thr Arg Arg Leu Arg Val  
2275 2280 2285

Ser His Ala Phe His Ser His Arg Met Asp Pro Met Leu Gly Glu Phe  
2290 2295 2300

Ala Gln Ile Ala Glu Ser Ala Glu Phe Gly Lys Pro Thr Thr Pro Leu  
305 2310 2315 2320

Val Ser Thr Leu Thr Gly Glu Leu Asp Arg Ala Ala Glu Met Ser Thr  
2325 2330 2335

Pro Gly Tyr Trp Val Arg Gln Ala Arg Glu Pro Val Arg Phe Ala Asp  
2340 2345 2350

Gly Val Gln Ala Leu Ala Ala Gln Gly Ile Gly Thr Val Val Glu Leu  
2355 2360 2365

Gly Pro Asp Gly Thr Leu Ala Ala Leu Val Arg Glu Cys Ala Thr Glu  
2370 2375 2380

Ser Asp Arg Val Gly Arg Ile Ser Ser Ile Pro Leu Met Arg Arg Glu  
385 2390 2395 2400

Arg Asp Glu Thr Arg Ser Val Met Thr Ala Leu Ala His Leu His Thr  
2405 2410 2415

Arg Gly Gly Glu Val Asp Trp Gln Ala Phe Phe Ala Gly Thr Gly Ala  
2420 2425 2430

Arg Gln Leu Glu Leu Pro Thr Tyr Ala Phe Gln Arg Gln His Tyr Trp  
2435 2440 2445

Ile Glu Ser Ser Ala Arg Pro Ala Arg Asp Arg Ala Asp Ile Gly Glu

2450

2455

2460

Val Ala Glu Gln Phe Trp Thr Ala Val Asp Gln Gly Asp Leu Ala Thr  
 465 2470 2475 2480

Leu Val Ala Ala Leu Asp Leu Gly Ala Asp Asp Thr Cys Ala Ser  
 2485 2490 2495

Leu Ser Asp Val Leu Pro Ala Leu Ser Ser Trp Arg Ser Gly Leu Arg  
 2500 2505 2510

Asn Arg Ser Leu Val Asp Ser Cys Arg Tyr Arg Ile Ser Trp His Ser  
 2515 2520 2525

Ser Arg Glu Val Pro Ala Pro Lys Ile Ser Gly Thr Trp Leu Leu Val  
 2530 2535 2540

Val Pro Gly Ala Ala Asp Asp Gly Leu Val Thr Ala Leu Thr Ser Ser  
 2545 2550 2555 2560

Leu Val Gly Gly Gly Ala Glu Val Val Arg Ile Gly Leu Ser Glu Glu  
 2565 2570 2575

Asp Pro His Arg Glu Asp Val Ala Gln Arg Leu Ala Asn Ala Leu Thr  
 2580 2585 2590

Asp Ala Gly Gln Leu Gly Gly Val Leu Ser Leu Leu Gly Leu Asp Glu  
 2595 2600 2605

Ser Pro Ala Pro Gly Phe Ser Cys Leu Pro Thr Gly Phe Ala Leu Thr  
 2610 2615 2620

Val Gln Leu Leu Arg Ala Leu Arg Lys Ala Asp Val Glu Ala Pro Phe  
 2625 2630 2635 2640

Trp Ala Val Thr Arg Gly Gly Val Ala Leu Glu Asp Val Arg Val Ser  
 2645 2650 2655

Pro Glu Gln Ala Leu Val Trp Gly Leu Leu Arg Val Ala Gly Leu Glu  
 2660 2665 2670

His Pro Glu Phe Trp Gly Gly Leu Ile Asp Leu Pro Ser Asp Trp Asp  
 2675 2680 2685

Asp Arg Leu Gly Ala Arg Leu Ala Gly Val Leu Ala Asp Gly Gly Glu  
 2690 2695 2700

Asp Gln Val Ala Ile Arg Arg Gly Gly Val Phe Val Arg Arg Leu Glu

705

2710

2715

2720

Arg Ala Gly Ala Ser Gly Ala Gly Ser Val Trp Arg Pro Arg Gly Thr  
 2725 2730 2735

Val Leu Val Thr Gly Gly Thr Gly Gly Leu Gly Ala His Val Ala Arg  
 2740 2745 2750

Trp Leu Ala Gly Ala Gly Ala Glu His Val Val Leu Thr Ser Arg Arg  
 2755 2760 2765

Gly Ala Asp Ala Pro Gly Ala Gly Glu Leu Arg Ala Glu Leu Glu Ala  
 2770 2775 2780

Leu Gly Ala Arg Val Ser Ile Val Pro Cys Asp Val Ala Asp Arg Asp  
 785 2790 2795 2800

Ala Val Ala Gly Val Leu Ala Gly Ile Gly Gly Glu Cys Pro Leu Thr  
 2805 2810 2815

Ala Val Val His Ala Ala Gly Val Gly Glu Ala Gly Asp Val Val Glu  
 2820 2825 2830

Met Gly Leu Ala Asp Phe Ala Ala Val Leu Ser Ala Lys Val Arg Gly  
 2835 2840 2845

Ala Ala Asn Leu Asp Glu Leu Leu Ala Asp Ser Glu Leu Asp Ala Phe  
 2850 2855 2860

Val Met Phe Ser Ser Val Ser Gly Val Trp Gly Ala Gly Gly Gln Gly  
 865 2870 2875 2880

Ala Tyr Ala Ala Ala Asn Ala Tyr Leu Asp Ala Leu Ala Glu Gln Arg  
 2885 2890 2895

Arg Ala Arg Gly Leu Val Gly Thr Ala Val Ala Trp Gly Pro Trp Ala  
 2900 2905 2910

Gly Asp Gly Met Ala Ala Gly Glu Thr Gly Ala Gln Leu His Arg Met  
 2915 2920 2925

Gly Leu Ala Ser Met Glu Pro Ser Ala Ala Leu Leu Ala Leu Gln Gly  
 2930 2935 2940

Ala Leu Asp Arg Asp Glu Thr Ser Leu Val Val Ala Asp Val Asp Trp  
 945 2950 2955 2960

Ala Arg Phe Ala Pro Ala Phe Thr Ser Ala Arg Arg Arg Pro Leu Leu

Asp Thr Ile Asp Glu Ala Arg Ala Ala Leu Glu Thr Thr Gly Glu Gln  
2980 2985 2990

Ala Gly Thr Gly Lys Pro Val Glu Leu Thr Gln Arg Leu Ala Gly Leu  
2995 3000 3005

Ser Arg Lys Glu Arg Asp Asp Ala Val Leu Asp Leu Val Arg Ala Glu  
3010 3015 3020

Thr Ala Ala Val Leu Gly Arg Asp Asp Ala Thr Ala Leu Ala Pro Ser  
025 3030 3035 3040

Arg Pro Phe Gln Glu Leu Gly Phe Asp Ser Leu Met Ala Val Glu Leu  
3045 3050 3055

Arg Asn Arg Leu Asn Thr Ala Thr Gly Ile Gln Leu Pro Ala Ser Thr  
3060 3065 3070

Ile Phe Asp Tyr Pro Asn Ala Glu Ser Leu Ser Arg His Leu Cys Ala  
3075 3080 3085

Glu Leu Phe Pro Thr Glu Thr Thr Val Asp Ser Ala Leu Ala Glu Leu  
3090 3095 3100

Asp Arg Ile Glu Gln Gln Leu Ser Met Leu Thr Gly Glu Ala Arg Ala  
105 3110 3115 3120

Arg Asp Arg Ile Ala Thr Arg Leu Arg Ala Leu His Glu Lys Trp Asn  
3125 3130 3135

Ser Ala Ala Glu Val Pro Thr Gly Ala Asp Val Leu Ser Thr Leu Asp  
3140 3145 3150

Ser Ala Thr His Asp Glu Ile Phe Glu Phe Ile Asp Asn Glu Leu Asp  
3155 3160 3165

Leu Ser  
3170

<210> 47

<211> 14775

<212> DNA

<213> Saccharopolyspora spinosa

<220>

<221> CDS

<222> (1)..(14772)

<223> ORF21; Polyketidsynthase

<400> 47

atg gcc aat gaa gaa aag ctc ttc ggc tat ctg aag aag gta act gcg	48
Met Ala Asn Glu Glu Lys Leu Phe Gly Tyr Leu Lys Lys Val Thr Ala	
1 5 10 15	
gac ctg cat cag acc cgg cag cgc ctg ctc gcg gcc gag agc cgg agt	96
Asp Leu His Gln Thr Arg Gln Arg Leu Leu Ala Ala Glu Ser Arg Ser	
20 25 30	
cag gag ccg atc gcg atc gtc tcg gcg agc tgc cga ctg ccc ggc ggc	144
Gln Glu Pro Ile Ala Ile Val Ser Ala Ser Cys Arg Leu Pro Gly Gly	
35 40 45	
gtc gac tct ccc gaa gcg ctc tgg caa ctc gtg cgc act ggc acc gac	192
Val Asp Ser Pro Glu Ala Leu Trp Gln Leu Val Arg Thr Gly Thr Asp	
50 55 60	
gcc atc tcg gag ttc ccc gcc gac cgg ggc tgg gat ctc ggc cgg ttg	240
Ala Ile Ser Glu Phe Pro Ala Asp Arg Gly Trp Asp Leu Gly Arg Leu	
65 70 75 80	
tac gat ccc gac ccg aac cac cag gga acg tcg tac acg cgg gcc ggc	288
Tyr Asp Pro Asp Pro Asn His Gln Gly Thr Ser Tyr Thr Arg Ala Gly	
85 90 95	
ggt ttc ctc gca gga gcg ggc gat ttc gac ccc gcc atg ttc ggg att	336
Gly Phe Leu Ala Gly Ala Gly Asp Phe Asp Pro Ala Met Phe Gly Ile	
100 105 110	
tcg ccg cgt gag gcg ttg gcg atg gac ccg cag caa cgg ttg ttg ctg	384
Ser Pro Arg Glu Ala Leu Ala Met Asp Pro Gln Gln Arg Leu Leu Leu	
115 120 125	
gag ctg tcc tgg gag gcc ctc gaa cgg gcg ggc ata gac ccg aca tcc	432
Glu Leu Ser Trp Glu Ala Leu Glu Arg Ala Gly Ile Asp Pro Thr Ser	
130 135 140	
ctg cgc ggc agc aag acc ggt gtc ttc ggt ggt gtc acg ccc cag gag	480
Leu Arg Gly Ser Lys Thr Gly Val Phe Gly Gly Val Thr Pro Gln Glu	
145 150 155 160	
tac ggg ccg tcc ttg cag gag atg agc cga aac gct ggg ggt ttt gga	528
Tyr Gly Pro Ser Leu Gln Glu Met Ser Arg Asn Ala Gly Gly Phe Gly	

165

170

175

ctc acc ggg cgg atg gtg agt gtg gcg tcg ggt cgg gtt gcg tat tcg 576  
 Leu Thr Gly Arg Met Val Ser Val Ala Ser Gly Arg Val Ala Tyr Ser  
           180                          185                          190

ttt ggt ttt gag ggt cct gcg gtg acg gtg gat acg gcg tgt tcg tcg 624  
 Phe Gly Phe Glu Gly Pro Ala Val Thr Val Asp Thr Ala Cys Ser Ser  
           195                          200                          205

tcg ttg gtg gcc ctg cat ttg gcg tgt cag tcg ttg cgt tcc ggc gaa 672  
 Ser Leu Val Ala Leu His Leu Ala Cys Gln Ser Leu Arg Ser Gly Glu  
           210                          215                          220

tgc gat ctc gcg ctg gcc ggc ggt gtg acg gtg atg gcg aca ccg gcg 720  
 Cys Asp Leu Ala Leu Ala Gly Gly Val Thr Val Met Ala Thr Pro Ala  
           225                          230                          235                          240

acg ttc gtg gag ttc tcc cgt cag cgt ggt ttg gct ccg gac ggg cgg 768  
 Thr Phe Val Glu Phe Ser Arg Gln Arg Gly Leu Ala Pro Asp Gly Arg  
                           245                          250                          255

tgc aag tcg ttc gcg gct gcc gcg gat ggc acc ggg tgg ggt gag ggt 816  
 Cys Lys Ser Phe Ala Ala Ala Ala Asp Gly Thr Gly Trp Gly Glu Gly  
                           260                          265                          270

gcc ggt ctg gtg ttg ctg gag cgg ttg tcg gat gcg cgg cgg aat ggg 864  
 Ala Gly Leu Val Leu Leu Glu Arg Leu Ser Asp Ala Arg Arg Asn Gly  
           275                          280                          285

cac gag gtt ctg gcg gtg gtg cgg ggt agc gcg gtg aac cag gac ggc 912  
 His Glu Val Leu Ala Val Val Arg Gly Ser Ala Val Asn Gln Asp Gly  
           290                          295                          300

gcg tcg aat ggt ttg act gcg ccg aat ggt ccg tcg cag cag cgg gtg 960  
 Ala Ser Asn Gly Leu Thr Ala Pro Asn Gly Pro Ser Gln Gln Arg Val  
           305                          310                          315                          320

atc acc cag gcg ttg gcg agt gcg ggg ctg tcg gtt tcc gat gtg gat 1008  
 Ile Thr Gln Ala Leu Ala Ser Ala Gly Leu Ser Val Ser Asp Val Asp  
                           325                          330                          335

gcg gtc gag gca cat ggg acc ggg acc acg ttg ggt gat ccg atc gag 1056  
 Ala Val Glu Ala His Gly Thr Gly Thr Thr Leu Gly Asp Pro Ile Glu  
           340                          345                          350

gca cag gcc ctg atc gcc acg tac ggg cag ggc cgg gag aag gat cgg 1104  
 Ala Gln Ala Leu Ile Ala Thr Tyr Gly Gln Gly Arg Glu Lys Asp Arg

355

360

365

ccg ttg tgg ttg ggg tcg gtc aag tcc aac atc ggt cac acg cag gcg	1152
Pro Leu Trp Leu Gly Ser Val Lys Ser Asn Ile Gly His Thr Gln Ala	
370 375 380	
gcc gct ggc gtt gcc ggc gtc atc aag atg gtc ttg gcg atg cgg cac	1200
Ala Ala Gly Val Ala Gly Val Ile Lys Met Val Leu Ala Met Arg His	
385 390 395 400	
ggg cag ctg ccc gcc acg ttg cat gtg gat gag ccc acg tcg gcg gtg	1248
Gly Gln Leu Pro Ala Thr Leu His Val Asp Glu Pro Thr Ser Ala Val	
405 410 415	
gac tgg tcg gcg ggt tcg gtc cgg ctt ctc acg gag aac acg ccc tgg	1296
Asp Trp Ser Ala Gly Ser Val Arg Leu Leu Thr Glu Asn Thr Pro Trp	
420 425 430	
ccg gac agt ggt cgt cct tgc cgg gtg ggg gtg tcg tcg ttc ggg atc	1344
Pro Asp Ser Gly Arg Pro Cys Arg Val Gly Val Ser Ser Phe Gly Ile	
435 440 445	
agc ggc acc aac gca cat gtg att ctc gaa cag tct cca gtc gag cag	1392
Ser Gly Thr Asn Ala His Val Ile Leu Glu Gln Ser Pro Val Glu Gln	
450 455 460	
ggc gaa ccg gcc ggg ccg gtc gaa ggc gag cgg gaa ccg gat gta gcc	1440
Gly Glu Pro Ala Gly Pro Val Glu Gly Glu Arg Glu Pro Asp Val Ala	
465 470 475 480	
gtc ccc gtg gtg cct tgg gtg ctg tcg ggt aag aca ccg gag gct gcg	1488
Val Pro Val Val Pro Trp Val Leu Ser Gly Lys Thr Pro Glu Ala Ala	
485 490 495	
cgg gcg cag gcc gaa ccg gtg cat tcg cat atc gag gac cgg ccg ggg	1536
Arg Ala Gln Ala Glu Arg Val His Ser His Ile Glu Asp Arg Pro Gly	
500 505 510	
ctg tcg ccg gtg gat gtg gcg tat tcg cta gga atg aca cgc gcg gcg	1584
Leu Ser Pro Val Asp Val Ala Tyr Ser Leu Gly Met Thr Arg Ala Ala	
515 520 525	
ctg gat gaa cgc gca gtg gtg ttg ggc tcg gac cgt gcc gcg ctc ctg	1632
Leu Asp Glu Arg Ala Val Val Leu Gly Ser Asp Arg Ala Ala Leu Leu	
530 535 540	
acc ggg ttg agg gca ttc gcc gac ggc tgc gat gcg ccc gaa gtg gtt	1680
Thr Gly Leu Arg Ala Phe Ala Asp Gly Cys Asp Ala Pro Glu Val Val	

545	550	555	560	
tcg ggg tct gtg ggg ctt ggt ggc cgc gtc ggg ttc gtg ttc tcg ggt	1728			
Ser Gly Ser Val Gly Leu Gly Gly Arg Val Gly Phe Val Phe Ser Gly				
565	570	575		
cag ggt ggt cag tgg ccg ggg atg ggc cgg ggg ctc tac tcg gtg ttt	1776			
Gln Gly Gly Gln Trp Pro Gly Met Gly Arg Gly Leu Tyr Ser Val Phe				
580	585	590		
ccg gtg ttc gcc gac gcg ttc gac gag gct tgc gcg gag ttg gat gca	1824			
Pro Val Phe Ala Asp Ala Phe Asp Glu Ala Cys Ala Glu Leu Asp Ala				
595	600	605		
cac ctg ggc cag gaa ctg cgg gtt cgg gat gtg gtg ttc ggt tcg caa	1872			
His Leu Gly Gln Glu Leu Arg Val Arg Asp Val Val Phe Gly Ser Gln				
610	615	620		
gcg tgg ttg ctg gat cgg acg gtg tgg gcg cag tcg ggt ttg ttc gcg	1920			
Ala Trp Leu Leu Asp Arg Thr Val Trp Ala Gln Ser Gly Leu Phe Ala				
625	630	635	640	
ttg cag att ggc ttg ctg cgg ctg ctg ggt tcg tgg ggt gtt cgg ccg	1968			
Leu Gln Ile Gly Leu Leu Arg Leu Leu Gly Ser Trp Gly Val Arg Pro				
645	650	655		
gat gtg gtg ttg ggg cac tcg gtg ggt gag ctg gct gcg gtg cat gcg	2016			
Asp Val Val Leu Gly His Ser Val Gly Glu Leu Ala Ala Val His Ala				
660	665	670		
gct ggt gtg ttg tcg ttg tcg gag gcc gcg cgg ttg gtg gcg ggt cgc	2064			
Ala Gly Val Leu Ser Leu Ser Glu Ala Ala Arg Leu Val Ala Gly Arg				
675	680	685		
gcc cgg ttg atg cag gcg ttg cct tct ggt ggt gcc atg ctc gcg gtc	2112			
Ala Arg Leu Met Gln Ala Leu Pro Ser Gly Gly Ala Met Leu Ala Val				
690	695	700		
gct acg ggt gag ttt cag gtc gat cct ctg ctg gat ggg gtg cgg gac	2160			
Ala Thr Gly Glu Phe Gln Val Asp Pro Leu Leu Asp Gly Val Arg Asp				
705	710	715	720	
cgg atc ggt atc gcg gcg gtg aat ggc ccg gaa tcg gtt gtg ctc tct	2208			
Arg Ile Gly Ile Ala Ala Val Asn Gly Pro Glu Ser Val Val Leu Ser				
725	730	735		
ggt gac cgc gag ctg ctc acc gag atc gct gat cgg ttg cac gat cag	2256			
Gly Asp Arg Glu Leu Leu Thr Glu Ile Ala Asp Arg Leu His Asp Gln				



740	745	750	
ggg tgc cgg acc cgg tgg ttg cgg gtg tgc cat gct ttc cat tgc ccc			2304
Gly Cys Arg Thr Arg Trp Leu Arg Val Ser His Ala Phe His Ser Pro			
755	760	765	
cat atg gag ccg atg ctg gag gag ttc gcc cag atc tcc cga ggc cgc			2352
His Met Glu Pro Met Leu Glu Glu Phe Ala Gln Ile Ser Arg Gly Arg			
770	775	780	
gaa tat cac gca ccg gaa ctg ccg atc atc tgc acc ctg atc ggt gag			2400
Glu Tyr His Ala Pro Glu Leu Pro Ile Ile Ser Thr Leu Ile Gly Glu			
785	790	795	800
ctg gac ggt ggt cga gtg atg ggc act ccc gag tac tgg gtg cgt cag			2448
Leu Asp Gly Gly Arg Val Met Gly Thr Pro Glu Tyr Trp Val Arg Gln			
	805	810	815
gtg cgt gag ccc gtc cgt ttc gcc gag ggt gtc cag gcg ctt gtc ggt			2496
Val Arg Glu Pro Val Arg Phe Ala Glu Gly Val Gln Ala Leu Val Gly			
	820	825	830
cag ggt gtc ggc acg att gtc gaa ttg ggt ccg gac ggg gcg ttg tgc			2544
Gln Gly Val Gly Thr Ile Val Glu Leu Gly Pro Asp Gly Ala Leu Ser			
	835	840	845
acg ttg gtc gag gag tgt gtg gcg gaa tcc ggg cgg gtg gcc ggg atc			2592
Thr Leu Val Glu Glu Cys Val Ala Glu Ser Gly Arg Val Ala Gly Ile			
	850	855	860
ccg ctg atg cgc aag gac cgc gac gag gcg cga acc gtg ctg gca gct			2640
Pro Leu Met Arg Lys Asp Arg Asp Glu Ala Arg Thr Val Leu Ala Ala			
	865	870	875
ttg gcg cag atc cac acc cgt ggt ggt gag gtg gac tgg cgg tgc ttt			2688
Leu Ala Gln Ile His Thr Arg Gly Gly Glu Val Asp Trp Arg Ser Phe			
	885	890	895
ttc gcc ggt acc ggg gcg aag caa gtc gac ctg ccc acc tac gcc ttc			2736
Phe Ala Gly Thr Gly Ala Lys Gln Val Asp Leu Pro Thr Tyr Ala Phe			
	900	905	910
cag cgg cag cgg tac tgg ctg gca tcc acc ggg cgt gcg ggt gac gtg			2784
Gln Arg Gln Arg Tyr Trp Leu Ala Ser Thr Gly Arg Ala Gly Asp Val			
	915	920	925
acc gcc gcc gga ttg gcc gag gcg gac cat ccg ctg ctc ggt gcg gtg			2832
Thr Ala Ala Gly Leu Ala Glu Ala Asp His Pro Leu Leu Gly Ala Val			

930

935

940

gtt gcg ttg gca gac ggc gaa ggt gtg gtg ctg acc ggt cgg ttg aca 2880  
 Val Ala Leu Ala Asp Gly Glu Gly Val Val Leu Thr Gly Arg Leu Thr  
 945 950 955 960

gcg ggt tcg cat ccg tgg ttg tcc gat cac cgg gtg ctg ggc gaa atc 2928  
 Ala Gly Ser His Pro Trp Leu Ser Asp His Arg Val Leu Gly Glu Ile  
 965 970 975

gtc gtc ccc ggc acc gcg atc gtc gag ctg gtg tgg cac gtc ggc gag 2976  
 Val Val Pro Gly Thr Ala Ile Val Glu Leu Val Trp His Val Gly Glu  
 980 985 990

cgc ctc ggt tgt ggc cgg gtg gaa gaa ctg gct ttg gaa gcg ccc ctg 3024  
 Arg Leu Gly Cys Gly Arg Val Glu Glu Leu Ala Leu Glu Ala Pro Leu  
 995 1000 1005

atc ctg ccg gat cat gga gcg gtc cag gtt cag gtg ctg gtg gga ccg 3072  
 Ile Leu Pro Asp His Gly Ala Val Gln Val Gln Val Leu Val Gly Pro  
 1010 1015 1020

ccc ggg gaa tcc gga gcc cgg tcg gtg gcg ctc tac tcc tgt cct ggc 3120  
 Pro Gly Glu Ser Gly Ala Arg Ser Val Ala Leu Tyr Ser Cys Pro Gly  
 1025 1030 1035 1040

gag gcg atc gaa ccc gag tgg aag aag cac gcg acg ggc gtg ctt ctc 3168  
 Glu Ala Ile Glu Pro Glu Trp Lys Lys His Ala Thr Gly Val Leu Leu  
 1045 1050 1055

cca ccc gtg gcc gcc gag aac cat gag ctg acc gca tgg ccc ccg gag 3216  
 Pro Pro Val Ala Ala Glu Asn His Glu Leu Thr Ala Trp Pro Pro Glu  
 1060 1065 1070

aat gcg acc gaa atc gat gca gac ggg gtc tac gca ttc ctt gaa ggg 3264  
 Asn Ala Thr Glu Ile Asp Ala Asp Gly Val Tyr Ala Phe Leu Glu Gly  
 1075 1080 1085

cac ggt ttc gcg tac gga ccg gcc ttt aga tgt ctg cgc ggt gcc tgg 3312  
 His Gly Phe Ala Tyr Gly Pro Ala Phe Arg Cys Leu Arg Gly Ala Trp  
 1090 1095 1100

cga cga ggc ggg gag gtg ttc gcc gaa gtc gca ttg ccg gat gac atg 3360  
 Arg Arg Gly Gly Glu Val Phe Ala Glu Val Ala Leu Pro Asp Asp Met  
 1105 1110 1115 1120

cag gcg ggg gtc gat cga ttc ggc gtc cac ccc gcg ttg ctg gac gcg 3408  
 Gln Ala Gly Val Asp Arg Phe Gly Val His Pro Ala Leu Leu Asp Ala

1125	1130	1135	
gtt ctg cat gcc gcc gca gcc gag acg tcg gtg gtc cag agc gaa gcg			3456
Val Leu His Ala Ala Ala Ala Glu Thr Ser Val Val Gln Ser Glu Ala			
1140	1145	1150	
cgg gtg ccg ttc tcg tgg cgt ggg gtg gaa ctt cgc gcc act gaa agc			3504
Arg Val Pro Phe Ser Trp Arg Gly Val Glu Leu Arg Ala Thr Glu Ser			
1155	1160	1165	
gcg gtg gtg ccg gcg cgc ctc tcg ttg act tcg gat gac gaa ctg tcg			3552
Ala Val Val Arg Ala Arg Leu Ser Leu Thr Ser Asp Asp Glu Leu Ser			
1170	1175	1180	
ttg gtc gca gtg gac ccg gct ggc cga ttc gtg gcc acg gtt gat tcg			3600
Leu Val Ala Val Asp Pro Ala Gly Arg Phe Val Ala Thr Val Asp Ser			
1185	1190	1195	1200
ctg gtg acc cga ccg atc tcc cgg cag cag gtg agg tct ggc gcg atc			3648
Leu Val Thr Arg Pro Ile Ser Arg Gln Gln Val Arg Ser Gly Ala Ile			
1205	1210	1215	
ggt gat tgc ctg ttc gag gtg gag tgg cac cgg aag gcg ttg ttg gga			3696
Gly Asp Cys Leu Phe Glu Val Glu Trp His Arg Lys Ala Leu Leu Gly			
1220	1225	1230	
aca acc gcc ggc gac gac ctt gcc atc gtc ggt gac ggt ccc agt tgg			3744
Thr Thr Ala Gly Asp Asp Leu Ala Ile Val Gly Asp Gly Pro Ser Trp			
1235	1240	1245	
ccg gaa tcg gtg cgc gca acc gca cgg ttc gcg acc ctg gat gag ttc			3792
Pro Glu Ser Val Arg Ala Thr Ala Arg Phe Ala Thr Leu Asp Glu Phe			
1250	1255	1260	
cgt gcg gcc gtg gac tcg gac gtt cct gcc ccg ggt tcg gtg ttg gtc			3840
Arg Ala Ala Val Asp Ser Asp Val Pro Ala Pro Gly Ser Val Leu Val			
1265	1270	1275	1280
gca gct atg tcg gcc gaa gag gtc gag ggt gga tcc ctg ccg tcg cgc			3888
Ala Ala Met Ser Ala Glu Glu Val Glu Gly Gly Ser Leu Pro Ser Arg			
1285	1290	1295	
gcc caa gag tcg acc tcc gat ctg ctg gct ctc gtg cag tcg tgg ctt			3936
Ala Gln Glu Ser Thr Ser Asp Leu Leu Ala Leu Val Gln Ser Trp Leu			
1300	1305	1310	
gcg gac gag ccg ttc gcc gaa tcc cag ctc gtg gtc gtc acg cgt gca			3984
Ala Asp Glu Arg Phe Ala Glu Ser Gln Leu Val Val Val Thr Arg Ala			

1315	1320	1325	
gcg gtg tcg gcc gac tcg gat tcg gac gtc gcg gac ctg gtg ggt gcg	4032		
Ala Val Ser Ala Asp Ser Asp Ser Asp Val Ala Asp Leu Val Gly Ala			
1330	1335	1340	
tcg tcg tgg ggg ttg ttg agt tca gcc cag tcg gag aac ccg ggt cgc	4080		
Ser Ser Trp Gly Leu Leu Ser Ser Ala Gln Ser Glu Asn Pro Gly Arg			
1345	1350	1355	1360
ttc gtg ctg gtg gac gtg gac ggc aca cct gag tcg tgg cag gcg ttg	4128		
Phe Val Leu Val Asp Val Asp Gly Thr Pro Glu Ser Trp Gln Ala Leu			
1365	1370	1375	
ccg gcc gcc gtg cga gca gga gaa ccg cag ctg gca ctt cgg cgc ggc	4176		
Pro Ala Ala Val Arg Ala Gly Glu Pro Gln Leu Ala Leu Arg Arg Gly			
1380	1385	1390	
gtg gcg ctg gtg cct cgg ttg gcg cga ctc acg gtg cgc gag gag ggc	4224		
Val Ala Leu Val Pro Arg Leu Ala Arg Leu Thr Val Arg Glu Glu Gly			
1395	1400	1405	
tcc tcc ccg caa ctc gac acg gac ggg acc gtc ctc atc acg ggt ggc	4272		
Ser Ser Pro Gln Leu Asp Thr Asp Gly Thr Val Leu Ile Thr Gly Gly			
1410	1415	1420	
acc ggt gcg ttg ggg gga gtg gtt gcc cgt cac ctg gtg gag gag cac	4320		
Thr Gly Ala Leu Gly Gly Val Val Ala Arg His Leu Val Glu Glu His			
1425	1430	1435	1440
ggg att cgg cgt ttg gtg ttg gca ggc cgg cgt ggc tgg aat gcg cct	4368		
Gly Ile Arg Arg Leu Val Leu Ala Gly Arg Arg Gly Trp Asn Ala Pro			
1445	1450	1455	
gga gtc cac gag ttg gtg gat gag ctg gcg cgc gcg ggc gcc gtg gtt	4416		
Gly Val His Glu Leu Val Asp Glu Leu Ala Arg Ala Gly Ala Val Val			
1460	1465	1470	
gag gtg gtg gct tgc gat gtg gct gac cgc acc gat ctg gag cac gtg	4464		
Glu Val Val Ala Cys Asp Val Ala Asp Arg Thr Asp Leu Glu His Val			
1475	1480	1485	
ctg gcc gcc att ccg gtc gac tgg ccg ctg cgg ggg atc gtg cat acc	4512		
Leu Ala Ala Ile Pro Val Asp Trp Pro Leu Arg Gly Ile Val His Thr			
1490	1495	1500	
gct ggg gtg ctg gcc gac gga gtg atc ggg tcc ttg tcg gcg gcg gat	4560		
Ala Gly Val Leu Ala Asp Gly Val Ile Gly Ser Leu Ser Ala Ala Asp			

1505	1510	1515	1520	
gtg ggc acg gtg ttt gcc ccg aag gtg acg ggg gca tgg cat ctg cac				4608
Val Gly Thr Val Phe Ala Pro Lys Val Thr Gly Ala Trp His Leu His				
1525		1530	1535	
gag ttg acc cgc gat ctg gat ctg tcg ttc ttc gtt ctt ttc tct tcc				4656
Glu Leu Thr Arg Asp Leu Asp Leu Ser Phe Phe Val Leu Phe Ser Ser				
1540	1545		1550	
ttc tcc ggg att gcg ggt gcc gca ggg cag gcc aac tac gcg gcg gcg				4704
Phe Ser Gly Ile Ala Gly Ala Ala Gly Gln Ala Asn Tyr Ala Ala Ala				
1555	1560		1565	
aac acg ttc ctg gat gca ttg gcg cgt tat cgc cgg gcg cgt ggg ctg				4752
Asn Thr Phe Leu Asp Ala Leu Ala Arg Tyr Arg Arg Ala Arg Gly Leu				
1570	1575		1580	
cct ggg ttg tcg ttg gcg tgg gga ctg tgg gcg caa ccc agc ggt atg				4800
Pro Gly Leu Ser Leu Ala Trp Gly Leu Trp Ala Gln Pro Ser Gly Met				
1585	1590	1595	1600	
acg agt ggc ttg gac gcg gcg tcg gtg gag cgg ttg gcg cgg acg ggc				4848
Thr Ser Gly Leu Asp Ala Ala Ser Val Glu Arg Leu Ala Arg Thr Gly				
1605	1610		1615	
atc gca gaa ctt tcc acg gag gat gga ctc cgc ctg ttc gat gcc gcg				4896
Ile Ala Glu Leu Ser Thr Glu Asp Gly Leu Arg Leu Phe Asp Ala Ala				
1620	1625		1630	
ttc gcg aag gac cgg gct tgc gtc gtt gcc gct cga ttg gac agg gcg				4944
Phe Ala Lys Asp Arg Ala Cys Val Val Ala Ala Arg Leu Asp Arg Ala				
1635	1640		1645	
ctg ctg gtc ggg aac gga cga tcg cac gcg att ccg gcg ctg ttg agc				4992
Leu Leu Val Gly Asn Gly Arg Ser His Ala Ile Pro Ala Leu Leu Ser				
1650	1655		1660	
gcg ttg gtt cct gtt cgc ggc ggt gtg gcg agg aaa aca gcc aat tct				5040
Ala Leu Val Pro Val Arg Gly Gly Val Ala Arg Lys Thr Ala Asn Ser				
1665	1670	1675	1680	
cag gcc gcg gat gag gac gca ctg ttg ggt ttg gtg cgg gag cac gtt				5088
Gln Ala Ala Asp Glu Asp Ala Leu Leu Gly Leu Val Arg Glu His Val				
1685	1690		1695	
tcg gcc gtg ctg ggt tat tcg ggt gcg gtc gag gtt ggg ggc gac cgt				5136
Ser Ala Val Leu Gly Tyr Ser Gly Ala Val Glu Val Gly Gly Asp Arg				

1700

1705

1710

gct ttc cgt gat ctg ggt ttt gat tcg ttg tct ggc gtg gag ttg cgg 5184  
 Ala Phe Arg Asp Leu Gly Phe Asp Ser Leu Ser Gly Val Glu Leu Arg  
 1715 1720 1725

aac cgc ctt gcc ggg gtg ctg ggg gtg cgg ttg ccg gcg act gcg gtg 5232  
 Asn Arg Leu Ala Gly Val Leu Gly Val Arg Leu Pro Ala Thr Ala Val  
 1730 1735 1740

ttc gac tat ccg acg ccg cgg gcg ctg gcg cgt ttc ctg cat cag gaa 5280  
 Phe Asp Tyr Pro Thr Pro Arg Ala Leu Ala Arg Phe Leu His Gln Glu  
 1745 1750 1755 1760

ctg gca ggc gag gtc gcg tcc acg tcg acg ccg gtg acc agg gca gcg 5328  
 Leu Ala Gly Glu Val Ala Ser Thr Ser Thr Pro Val Thr Arg Ala Ala  
 1765 1770 1775

agt gcc gaa gag gat ctt gtt gcg att gtc ggg atg gga tgt cgt ttt 5376  
 Ser Ala Glu Glu Asp Leu Val Ala Ile Val Gly Met Gly Cys Arg Phe  
 1780 1785 1790

ccg ggt ggg gtg tcg tcg ccg gag gag ctt tgg cgg ctg gtg gcc ggc 5424  
 Pro Gly Gly Val Ser Ser Pro Glu Glu Leu Trp Arg Leu Val Ala Gly  
 1795 1800 1805

ggc gtg gat gcg gtg gct ggg ttc cca gac gat cgc ggc tgg gat ctc 5472  
 Gly Val Asp Ala Val Ala Gly Phe Pro Asp Asp Arg Gly Trp Asp Leu  
 1810 1815 1820

gcg gcg ttg tac gat cct gat ccc gat cgt ctc ggg acc tcg tat gtg 5520  
 Ala Ala Leu Tyr Asp Pro Asp Pro Asp Arg Leu Gly Thr Ser Tyr Val  
 1825 1830 1835 1840

tgt gag ggc ggg ttt ctg cgg gac gcg gcg gag ttc gat gct gac atg 5568  
 Cys Glu Gly Gly Phe Leu Arg Asp Ala Ala Glu Phe Asp Ala Asp Met  
 1845 1850 1855

ttc ggc atc agc ccg cgt gag gcg ttg gcg atg gat ccg cag cag cgg 5616  
 Phe Gly Ile Ser Pro Arg Glu Ala Leu Ala Met Asp Pro Gln Gln Arg  
 1860 1865 1870

ttg ctg ctg gag gtc gcc tgg gaa acc ttg gag cgg gct ggg atc gat 5664  
 Leu Leu Leu Glu Val Ala Trp Glu Thr Leu Glu Arg Ala Gly Ile Asp  
 1875 1880 1885

ccg ttc tcg ttg cac ggc agc cgg acc ggt gtg ttc gcg ggc ttg atg 5712  
 Pro Phe Ser Leu His Gly Ser Arg Thr Gly Val Phe Ala Gly Leu Met

1890	1895	1900	
tac cac gac tat ggg gcc cga ttc att acc aga gca ccg gag ggc ttc			5760
Tyr His Asp Tyr Gly Ala Arg Phe Ile Thr Arg Ala Pro Glu Gly Phe			
1905	1910	1915	1920
gaa ggg cac ctc ggg acg ggc aat gcg ggg agc gtg ctg tcg ggt cgg			5808
Glu Gly His Leu Gly Thr Gly Asn Ala Gly Ser Val Leu Ser Gly Arg			
1925	1930	1935	
gtt gcg tat tcg ttt ggt ttc gag ggt cct gcg gtg acg gtg gat acg			5856
Val Ala Tyr Ser Phe Gly Phe Glu Gly Pro Ala Val Thr Val Asp Thr			
1940	1945	1950	
gcg tgt tcg tcg tcg ttg gtg gcg tta cac ctg gcg ggt caa gca ctg			5904
Ala Cys Ser Ser Ser Leu Val Ala Leu His Leu Ala Gly Gln Ala Leu			
1955	1960	1965	
cgg gcc ggt gag tgc gaa ttc gcc ctt gcc ggt ggc gtc acg gtg atg			5952
Arg Ala Gly Glu Cys Glu Phe Ala Leu Ala Gly Gly Val Thr Val Met			
1970	1975	1980	
tcg acg ccg acg acg ttc gtg gag ttc tcc cgt caa ccg ggt ctg gct			6000
Ser Thr Pro Thr Thr Phe Val Glu Phe Ser Arg Gln Arg Gly Leu Ala			
1985	1990	1995	2000
ccg gat ggg ccg tgc aag tcg ttc gcg gcg gcc gcg gat ggc acc ggg			6048
Pro Asp Gly Arg Cys Lys Ser Phe Ala Ala Ala Ala Asp Gly Thr Gly			
2005	2010	2015	
tgg ggc gag ggt gcc ggt ctg gtg ttg ctg gag ccg ttg tcg gat gcc			6096
Trp Gly Glu Gly Ala Gly Leu Val Leu Leu Glu Arg Leu Ser Asp Ala			
2020	2025	2030	
cgg cgc aat ggg cac gag gtt ctg gcg gtg gtg ccg ggt agc gcg gtg			6144
Arg Arg Asn Gly His Glu Val Leu Ala Val Val Arg Gly Ser Ala Val			
2035	2040	2045	
aac cag gac ggc gcg tcg aat ggc ttg act gcg cca aat ggt ccg tca			6192
Asn Gln Asp Gly Ala Ser Asn Gly Leu Thr Ala Pro Asn Gly Pro Ser			
2050	2055	2060	
cag caa agg gtg atc acc cag gca ctc acg agt gcc ggg ctg tcc gtg			6240
Gln Gln Arg Val Ile Thr Gln Ala Leu Thr Ser Ala Gly Leu Ser Val			
2065	2070	2075	2080
tcc gac gtg gat gct gtg gag gcg cat ggg acg ggc acg ccg ctt ggt			6288
Ser Asp Val Asp Ala Val Glu Ala His Gly Thr Gly Thr Arg Leu Gly			

gat ccg atc gag gcg cag gcg ttg atc gct acg tac ggc cgg gat cgt Asp Pro Ile Glu Ala Gln Ala Leu Ile Ala Thr Tyr Gly Arg Asp Arg 2100 2105 2110	6336
gat ccc ggt cgg ccg ttg tgg ctg ggg tcg gtg aag tcg aat att ggt Asp Pro Gly Arg Pro Leu Trp Leu Gly Ser Val Lys Ser Asn Ile Gly 2115 2120 2125	6384
cac acc cag gcg gcg gcg ggt gtc gct ggt gtg atc aag atg gtg atg His Thr Gln Ala Ala Ala Gly Val Ala Gly Val Ile Lys Met Val Met 2130 2135 2140	6432
gcg atg cgg cag ggg gag ctg ccg cgc acg ttg cac gtg gac gag ccc Ala Met Arg Gln Gly Glu Leu Pro Arg Thr Leu His Val Asp Glu Pro 2145 2150 2155 2160	6480
tcc gcg cag gtg gac tgg tct gcg ggc acg gtc caa ctc ctc acg gag Ser Ala Gln Val Asp Trp Ser Ala Gly Thr Val Gln Leu Leu Thr Glu 2165 2170 2175	6528
aac acg ccc tgg ccc gac agc ggt cgt ctt cgc cgg gcg ggc gtg tca Asn Thr Pro Trp Pro Asp Ser Gly Arg Leu Arg Arg Ala Gly Val Ser 2180 2185 2190	6576
tcg ttc ggg atc agt ggc acc aac gcg cac ctg atc ctt gaa caa cct Ser Phe Gly Ile Ser Gly Thr Asn Ala His Leu Ile Leu Glu Gln Pro 2195 2200 2205	6624
ccg cga gag tcg cag cgc tca aca gag ccg gat tcg ggt tct gtc cgc Pro Arg Glu Ser Gln Arg Ser Thr Glu Pro Asp Ser Gly Ser Val Arg 2210 2215 2220	6672
gat ttt ccg gtg gtg ccg tgg atg gtg tcg ggc aaa aca ccc gaa gcg Asp Phe Pro Val Val Pro Trp Met Val Ser Gly Lys Thr Pro Glu Ala 2225 2230 2235 2240	6720
cta tcc gcc cag gca gat gca ttg atg tcc tac ttg agc aat cgc gtt Leu Ser Ala Gln Ala Asp Ala Leu Met Ser Tyr Leu Ser Asn Arg Val 2245 2250 2255	6768
gat gct tcc ccg cga gat atc ggt tat tcg ctt gcg gtg acc cgt ccg Asp Ala Ser Pro Arg Asp Ile Gly Tyr Ser Leu Ala Val Thr Arg Pro 2260 2265 2270	6816
gcg ttg gac cac cgc gct gtc gtg ctg ggt gcg gat cgt gcc gcg ttg Ala Leu Asp His Arg Ala Val Val Leu Gly Ala Asp Arg Ala Ala Leu	6864



2275	2280	2285	
ctg ccg ggc ttg aaa gcg ctg gcc gtt agt aat gac gct gcc gag gtg			6912
Leu Pro Gly Leu Lys Ala Leu Ala Val Ser Asn Asp Ala Ala Glu Val			
2290	2295	2300	
atc acc ggc act cgt gcc gct ggg ccg gtc gga ttc gtg ttc tcc ggt			6960
Ile Thr Gly Thr Arg Ala Ala Gly Pro Val Gly Phe Val Phe Ser Gly			
2305	2310	2315	2320
caa ggt ggt cag tgg ccc ggg atg gga agc ggg ctg cac tgg gcg ttt			7008
Gln Gly Gly Gln Trp Pro Gly Met Gly Ser Gly Leu His Ser Ala Phe			
2325	2330	2335	
ccg gtg ttc gcc gac gcg ttt gac gaa gcc tgc tgc gag ctg gat gcg			7056
Pro Val Phe Ala Asp Ala Phe Asp Glu Ala Cys Cys Glu Leu Asp Ala			
2340	2345	2350	
cat ctg ggg cag atg gcc cgg cta cga gat gtg ttg tcc ggt tgg gat			7104
His Leu Gly Gln Met Ala Arg Leu Arg Asp Val Leu Ser Gly Ser Asp			
2355	2360	2365	
acg caa ctt ctg gac cag acc ttg tgg gcg cag ccg ggc ctg ttc gcg			7152
Thr Gln Leu Leu Asp Gln Thr Leu Trp Ala Gln Pro Gly Leu Phe Ala			
2370	2375	2380	
ttg caa gtc gga ctg tgg gag ttg ttg ggt tgg tgg ggt gtc cgg ccc			7200
Leu Gln Val Gly Leu Trp Glu Leu Leu Gly Ser Trp Gly Val Arg Pro			
2385	2390	2395	2400
gct gtg gtg ctg ggc cac tgg gtc ggt gag ctg gcg gcg gcg ttc gcg			7248
Ala Val Val Leu Gly His Ser Val Gly Glu Leu Ala Ala Ala Phe Ala			
2405	2410	2415	
gct gga gtg ttg tgg ttg cgg gat gcg gct cgg ctg gtg gcg ggc cgt			7296
Ala Gly Val Leu Ser Leu Arg Asp Ala Ala Arg Leu Val Ala Gly Arg			
2420	2425	2430	
gcc cgg ttg atg caa gcc ctg cca act ggc ggt gcc atg ctg gct gcg			7344
Ala Arg Leu Met Gln Ala Leu Pro Thr Gly Gly Ala Met Leu Ala Ala			
2435	2440	2445	
gct gct gga gag gag cag ctg cgc ccg ttg ctg gcc gac tgc ggt gat			7392
Ala Ala Gly Glu Glu Gln Leu Arg Pro Leu Leu Ala Asp Cys Gly Asp			
2450	2455	2460	
cgt gtg ggg atc gcc gcg gtc aac gct ccc ggg tgg gtg gtg ctg tcc			7440
Arg Val Gly Ile Ala Ala Val Asn Ala Pro Gly Ser Val Val Leu Ser			

2465	2470	2475	2480	
ggt gat cgg gat gtg ctc gat gac att gcc ggt cgg ctg gac ggg caa				7488
Gly Asp Arg Asp Val Leu Asp Asp Ile Ala Gly Arg Leu Asp Gly Gln				
2485	2490	2495		
ggg atc cgg tcc agg tgg ttg cgg gtt tgc cat gcg ttt cat tgc cat				7536
Gly Ile Arg Ser Arg Trp Leu Arg Val Ser His Ala Phe His Ser His				
2500	2505	2510		
cgg atg gat ccg atg ctg gcg gag ttc acc gaa atc gcc cgg agc gtg				7584
Arg Met Asp Pro Met Leu Ala Glu Phe Thr Glu Ile Ala Arg Ser Val				
2515	2520	2525		
gac tac cgg tgc tca ggg ctg ccg atc gtg tgc acg ttg acg ggt gag				7632
Asp Tyr Arg Ser Ser Gly Leu Pro Ile Val Ser Thr Leu Thr Gly Glu				
2530	2535	2540		
ctc gat gag gtc ggc atg ccg gct acg ccg gag tat tgg gtg cgc cag				7680
Leu Asp Glu Val Gly Met Pro Ala Thr Pro Glu Tyr Trp Val Arg Gln				
2545	2550	2555	2560	
gtg cga gaa ccc gtc cgc ttc gcc gac ggt gtt gct gcg ctc gcg gct				7728
Val Arg Glu Pro Val Arg Phe Ala Asp Gly Val Ala Ala Leu Ala Ala				
2565	2570	2575		
cac ggt gtg agc acc gtc gtc gag gtc ggt ccg gat ggg gtg ttg tgc				7776
His Gly Val Ser Thr Val Val Glu Val Gly Pro Asp Gly Val Leu Ser				
2580	2585	2590		
gcg ctg gtg cag gag tgc gcg gcc gga tcc gat cag gcc gga cgg gtg				7824
Ala Leu Val Gln Glu Cys Ala Ala Gly Ser Asp Gln Gly Gly Arg Val				
2595	2600	2605		
gcc gcg gtt ccg ctc atg cgc agc aat cgc gac gag gcg cac acg gtg				7872
Ala Ala Val Pro Leu Met Arg Ser Asn Arg Asp Glu Ala His Thr Val				
2610	2615	2620		
aca acg gca ttg gcg cag atc cat gtg cgt ggt gct gag gtg gac tgg				7920
Thr Thr Ala Leu Ala Gln Ile His Val Arg Gly Ala Glu Val Asp Trp				
2625	2630	2635	2640	
cgg tgc ttt ttc gcc ggt acc ggg gca aag cag gtc gag ctg ccc acg				7968
Arg Ser Phe Phe Ala Gly Thr Gly Ala Lys Gln Val Glu Leu Pro Thr				
2645	2650	2655		
tat gcc ttc caa cga cag cgg tac tgg ctt gac tca cca tcc gaa ccg				8016
Tyr Ala Phe Gln Arg Gln Arg Tyr Trp Leu Asp Ser Pro Ser Glu Pro				

2660	2665	2670	
gtc ggg caa tcc gcc gat ccc gcg cgc cag tcg ggc ttc tgg gaa ctc			8064
Val Gly Gln Ser Ala Asp Pro Ala Arg Gln Ser Gly Phe Trp Glu Leu			
2675	2680	2685	
gtc gag cag gaa gat gtc agc gcg ctc agc gcc gct ctg cac att acc			8112
Val Glu Gln Glu Asp Val Ser Ala Leu Ser Ala Ala Leu His Ile Thr			
2690	2695	2700	
ggc gat cac gac gtg cag gcg tcc ctg gaa tcg gtg gtt ccg gtc ctc			8160
Gly Asp His Asp Val Gln Ala Ser Leu Glu Ser Val Val Pro Val Leu			
2705	2710	2715	2720
tcc tcc tgg cat cgc cgg atc cgc aac gaa tcc ctg gtg cac cag tgg			8208
Ser Ser Trp His Arg Arg Ile Arg Asn Glu Ser Leu Val His Gln Trp			
2725	2730	2735	
cgg tac cgg att tcc tgg cat gag cgg gca gat ttg cca gac ccc tcg			8256
Arg Tyr Arg Ile Ser Trp His Glu Arg Ala Asp Leu Pro Asp Pro Ser			
2740	2745	2750	
ttg tcg ggg aca tgg ctc gtc gtc gtg ccg gag ggg tgg tcg gcg agt			8304
Leu Ser Gly Thr Trp Leu Val Val Val Pro Glu Gly Trp Ser Ala Ser			
2755	2760	2765	
cgg caa gtt ctg cgt ttc aac gag atg ttc gag gaa cgg ggt tgc ccg			8352
Arg Gln Val Leu Arg Phe Asn Glu Met Phe Glu Glu Arg Gly Cys Pro			
2770	2775	2780	
gca gtt ctg ttc gag ctc gcc ggg cac gac gag gaa gcc ctg gcg caa			8400
Ala Val Leu Phe Glu Leu Ala Gly His Asp Glu Glu Ala Leu Ala Gln			
2785	2790	2795	2800
cga ttc cgc tcg ttg cct gtt gcg tca ggg gga ata agc ggc gtg ttg			8448
Arg Phe Arg Ser Leu Pro Val Ala Ser Gly Gly Ile Ser Gly Val Leu			
2805	2810	2815	
tcc ttg ctg gcg ctg gat gaa tcg ccg tcc tcg ccg aac gct gct ttg			8496
Ser Leu Leu Ala Leu Asp Glu Ser Pro Ser Ser Pro Asn Ala Ala Leu			
2820	2825	2830	
ccg aat ggc gcg ctg aac tcg ttg gta ctg ctg cga gct ctg cgg gcc			8544
Pro Asn Gly Ala Leu Asn Ser Leu Val Leu Leu Arg Ala Leu Arg Ala			
2835	2840	2845	
gcg gat gtg tcg gcg cca ttg tgg ttg gcg acg tgt ggt ggt gtc gcg			8592
Ala Asp Val Ser Ala Pro Leu Trp Leu Ala Thr Cys Gly Gly Val Ala			

2850	2855	2860	
gtc ggg gat gtg ccg gtg aac ccg ggg cag gcg ctg gtg tgg gga ctg			8640
Val Gly Asp Val Pro Val Asn Pro Gly Gln Ala Leu Val Trp Gly Leu			
2865	2870	2875	2880
ggt cgc gtc gtc ggt ctg gag cat ccg gcc tgg tgg ggt ggc ctg gtc			8688
Gly Arg Val Val Gly Leu Glu His Pro Ala Trp Trp Gly Gly Leu Val			
2885	2890	2895	
gac gtg ccg tgc ttg ctc gat gag gac gct cga gaa cgc ttg tcg gtc			8736
Asp Val Pro Cys Leu Leu Asp Glu Asp Ala Arg Glu Arg Leu Ser Val			
2900	2905	2910	
gtg ttg gca ggt ctt ggc gag gac gag atc gcg gta cgt ccc ggt ggt			8784
Val Leu Ala Gly Leu Gly Glu Asp Glu Ile Ala Val Arg Pro Gly Gly			
2915	2920	2925	
gtg ttc gtg cgg cgg ttg gaa cgc gct ggt gcg gcg tcg ggt gcc ggg			8832
Val Phe Val Arg Arg Leu Glu Arg Ala Gly Ala Ala Ser Gly Ala Gly			
2930	2935	2940	
tcg gtg tgg cgt cct cgg ggg acg gtg ttg gtg acg ggt ggt acg ggc			8880
Ser Val Trp Arg Pro Arg Gly Thr Val Leu Val Thr Gly Gly Thr Gly			
2945	2950	2955	2960
ggt ttg ggg gcg cat gtt gcc cgg tgg ttg gcg ggt gcc ggg gct gag			8928
Gly Leu Gly Ala His Val Ala Arg Trp Leu Ala Gly Ala Gly Ala Glu			
2965	2970	2975	
cat gtg gtg ttg acc agc cgt cga ggc gcg gcg gct ccg ggc gct gga			8976
His Val Val Leu Thr Ser Arg Arg Gly Ala Ala Ala Pro Gly Ala Gly			
2980	2985	2990	
gat ttg cgg gcg gag ctg gag gcg ctg ggc gct cgg gtt tcg atc acg			9024
Asp Leu Arg Ala Glu Leu Glu Ala Leu Gly Ala Arg Val Ser Ile Thr			
2995	3000	3005	
gcc tgc gac gtg gcc gat cgt gac gct ttg gcc gaa gtg ttg gcg acc			9072
Ala Cys Asp Val Ala Asp Arg Asp Ala Leu Ala Glu Val Leu Ala Thr			
3010	3015	3020	
att ccg gat gat tgc ccg ctg acc gcg gtg atg cat gcg gcg ggg gtc			9120
Ile Pro Asp Asp Cys Pro Leu Thr Ala Val Met His Ala Ala Gly Val			
3025	3030	3035	3040
ggt gaa gtc ggc gac gtg gcg tcg atg tgt ttg acc gac ttc gtt ggg			9168
Val Glu Val Gly Asp Val Ala Ser Met Cys Leu Thr Asp Phe Val Gly			

3045	3050	3055	
gtg ctg tcg gcg aag gca ggt ggt gcg gcg aat ctc gat gag ttg ctc			9216
Val Leu Ser Ala Lys Ala Gly Gly Ala Ala Asn Leu Asp Glu Leu Leu			
3060	3065	3070	
gcc gat gtc gag ctg gat gcc ttc gtg ctg ttc tca tcc gtc tcg ggt			9264
Ala Asp Val Glu Leu Asp Ala Phe Val Leu Phe Ser Ser Val Ser Gly			
3075	3080	3085	
gtg tgg ggt gct ggc ggg cag ggc gct tat gcg gcg gcg aat gcc tac			9312
Val Trp Gly Ala Gly Gly Gln Gly Ala Tyr Ala Ala Ala Asn Ala Tyr			
3090	3095	3100	
ttg gat gcg ttg gcg cag cag cgt cgg gca agg ggg ttg gtg ggg act			9360
Leu Asp Ala Leu Ala Gln Gln Arg Arg Ala Arg Gly Leu Val Gly Thr			
3105	3110	3115	3120
gcg gtt gcg tgg ggc ccg tgg gcc ggt gac gga atg gcc gca ggt gaa			9408
Ala Val Ala Trp Gly Pro Trp Ala Gly Asp Gly Met Ala Ala Gly Glu			
3125	3130	3135	
ggc ggt gca cag ctg cgc cgg gcc ggc ctg gtg cca atg gct gcg gat			9456
Gly Gly Ala Gln Leu Arg Arg Ala Gly Leu Val Pro Met Ala Ala Asp			
3140	3145	3150	
cgg gcg ttg ctg gca ctt cag ggc gca ttg gat cgt gac gag aca tcc			9504
Arg Ala Leu Leu Ala Leu Gln Gly Ala Leu Asp Arg Asp Glu Thr Ser			
3155	3160	3165	
ctg gtc gtg gcc gat atg gcg tgg gag agg ttc gcc ccg gtg ttc gcc			9552
Leu Val Val Ala Asp Met Ala Trp Glu Arg Phe Ala Pro Val Phe Ala			
3170	3175	3180	
atg tcc cgt cgg cgt ccg ctg ctc gac gag ctg ccc gaa gca cag cag			9600
Met Ser Arg Arg Arg Pro Leu Leu Asp Glu Leu Pro Glu Ala Gln Gln			
3185	3190	3195	3200
gcg ttg gcg gat gcg gag aac acc act gat gct gcg gac tcg gcc gtc			9648
Ala Leu Ala Asp Ala Glu Asn Thr Thr Asp Ala Ala Asp Ser Ala Val			
3205	3210	3215	
ccg cta ccg cgg ctc gcg ggc atg gca gcc gcc gaa cgc cgc cgc gcg			9696
Pro Leu Pro Arg Leu Ala Gly Met Ala Ala Ala Glu Arg Arg Arg Ala			
3220	3225	3230	
atg ctg gac ctg gtg ctg gcg gag gcc tcg att gtg ttg gga cac aac			9744
Met Leu Asp Leu Val Leu Ala Glu Ala Ser Ile Val Leu Gly His Asn			

3235

3240

3245

ggg tct gac cca gtt ggt ccc gac cgg gcg ttc cag gag ctc gga ttt 9792  
 Gly Ser Asp Pro Val Gly Pro Asp Arg Ala Phe Gln Glu Leu Gly Phe  
 3250 3255 3260

gat tcg ctg atg gcc gtc gaa ctg cgc aac agg ttg ggc gag gca aca 9840  
 Asp Ser Leu Met Ala Val Glu Leu Arg Asn Arg Leu Gly Glu Ala Thr  
 3265 3270 3275 3280

gga ttg agt ctg ccg gcc acg ttg atc ttc gat tat ccg agc cca tcc 9888  
 Gly Leu Ser Leu Pro Ala Thr Leu Ile Phe Asp Tyr Pro Ser Pro Ser  
 3285 3290 3295

gcg ctg gct gag cag ctg gtc ggc gag ctg gtg gga gcg cag ccc gcg 9936  
 Ala Leu Ala Glu Gln Leu Val Gly Glu Leu Val Gly Ala Gln Pro Ala  
 3300 3305 3310

acc acc gtc gtg gcc ggg gcc gat cca gtg gat gat ccg gtt gtc gtg 9984  
 Thr Thr Val Val Ala Gly Ala Asp Pro Val Asp Asp Pro Val Val Val  
 3315 3320 3325

gtc gcg atg gga tgc cgg tat ccg ggc gac gtc tgc tcg ccc gag gag 10032  
 Val Ala Met Gly Cys Arg Tyr Pro Gly Asp Val Cys Ser Pro Glu Glu  
 3330 3335 3340

ctg tgg cag ctg gtt tct gcg gga cgt gat gcg gta tcg acg ttc ccc 10080  
 Leu Trp Gln Leu Val Ser Ala Gly Arg Asp Ala Val Ser Thr Phe Pro  
 3345 3350 3355 3360

gtc gat cgg ggt tgg gac tgc aac acg ttg ttc gac ccg gat ccg gat 10128  
 Val Asp Arg Gly Trp Asp Cys Asn Thr Leu Phe Asp Pro Asp Pro Asp  
 3365 3370 3375

cgg gca ggc agt acc tat gtg cga gaa ggt gcc ttc ctg acc ggt gct 10176  
 Arg Ala Gly Ser Thr Tyr Val Arg Glu Gly Ala Phe Leu Thr Gly Ala  
 3380 3385 3390

gat cgg ttc gac gcc ggg ttc ttc ggc atc agc cct cgc gag gcg gcg 10224  
 Asp Arg Phe Asp Ala Gly Phe Phe Gly Ile Ser Pro Arg Glu Ala Arg  
 3395 3400 3405

gca atg gat ccg cag cag agg ttg ttg ctc gaa gtg gcg tgg gag gtt 10272  
 Ala Met Asp Pro Gln Gln Arg Leu Leu Leu Glu Val Ala Trp Glu Val  
 3410 3415 3420

ttc gaa cga gca gga atc gct ccg ctg tcg ttg cgg ggt agc agg acc 10320  
 Phe Glu Arg Ala Gly Ile Ala Pro Leu Ser Leu Arg Gly Ser Arg Thr

3425	3430	3435	3440	
ggt gtg ttc gcg ggg acc aat ggg cag gac cac ggt gcg aaa gtg gct				10368
Gly Val Phe Ala Gly Thr Asn Gly Gln Asp His Gly Ala Lys Val Ala				
3445		3450	3455	
gcc gcg ccg gag gcg gcg ggt cac ctc ctg acc gga aac gcc gcg agt				10416
Ala Ala Pro Glu Ala Ala Gly His Leu Leu Thr Gly Asn Ala Ala Ser				
3460	3465		3470	
gtc ctg gcc ggc cgg ctt tcc tac acg ttc ggc ctt gag ggg cct gcg				10464
Val Leu Ala Gly Arg Leu Ser Tyr Thr Phe Gly Leu Glu Gly Pro Ala				
3475	3480		3485	
gtg gcg gtg gat acc gcg tgt tcg tcg tcg ttg gtg gcg ttg cat ttg				10512
Val Ala Val Asp Thr Ala Cys Ser Ser Ser Leu Val Ala Leu His Leu				
3490	3495		3500	
gcg tgc cag tcg ctg cgt tcg ggt gag tgt gat atg gcg ttg gca ggt				10560
Ala Cys Gln Ser Leu Arg Ser Gly Glu Cys Asp Met Ala Leu Ala Gly				
3505	3510	3515	3520	
ggt gtg acg gtg atg tcg aca ccc ctg gct ttc ctc gag ttc tct cgt				10608
Gly Val Thr Val Met Ser Thr Pro Leu Ala Phe Leu Glu Phe Ser Arg				
3525	3530		3535	
cag cgc ggt ttg gcg cca gat ggt cgg tgc aag tcg ttt gcg gcc gct				10656
Gln Arg Gly Leu Ala Pro Asp Gly Arg Cys Lys Ser Phe Ala Ala Ala				
3540	3545		3550	
gcg gat ggc acc ggg tgg ggt gag ggt gcc ggc ctg gtg ttg ctg gag				10704
Ala Asp Gly Thr Gly Trp Gly Glu Gly Ala Gly Leu Val Leu Leu Glu				
3555	3560		3565	
cgg ttg tcg gat gct cgt cgg aat ggt cac cgg gtg ttg gcc gtg gtt				10752
Arg Leu Ser Asp Ala Arg Arg Asn Gly His Arg Val Leu Ala Val Val				
3570	3575		3580	
cgc ggg tct gcg gtg aat cag gat ggt gcg tcg aat ggc ctg act gcg				10800
Arg Gly Ser Ala Val Asn Gln Asp Gly Ala Ser Asn Gly Leu Thr Ala				
3585	3590	3595	3600	
ccg aat ggt ccg tcg cag cag cgg gtg att cgg cag gcc ctc gcg aat				10848
Pro Asn Gly Pro Ser Gln Gln Arg Val Ile Arg Gln Ala Leu Ala Asn				
3605	3610		3615	
gcg ggg ctg tcg gcg tcc gat gtg gat gtc gtg gag gcg cac ggg acc				10896
Ala Gly Leu Ser Ala Ser Asp Val Asp Val Val Glu Ala His Gly Thr				

3620

3625

3630

ggt acc ggg ctc ggg gat ccg atc gag gcg cag gcg ctg atc gcg aca 10944  
 Gly Thr Gly Leu Gly Asp Pro Ile Glu Ala Gln Ala Leu Ile Ala Thr  
 3635 3640 3645

tat ggg cag gag cgg gat cct gag cgg gcc ctg tgg ctg ggg tcg atc 10992  
 Tyr Gly Gln Glu Arg Asp Pro Glu Arg Ala Leu Trp Leu Gly Ser Ile  
 3650 3655 3660

aag tcc aac atc ggc cac acg cag gcg gcc ggt gtg gcg ggg gtc 11040  
 Lys Ser Asn Ile Gly His Thr Gln Ala Ala Ala Gly Val Ala Gly Val  
 3665 3670 3675 3680

atc aag atg gtg cag gcc atg cgg cac ggg gag ttg cct gcg acg ttg 11088  
 Ile Lys Met Val Gln Ala Met Arg His Gly Glu Leu Pro Ala Thr Leu  
 3685 3690 3695

cac gtg gac aag ccc act cca cag gtg gac tgg tct gcc ggg gcc gtt 11136  
 His Val Asp Lys Pro Thr Pro Gln Val Asp Trp Ser Ala Gly Ala Val  
 3700 3705 3710

cgg ctc ctc acc ggg aac acg ccc tgg ccc gag agc ggc cgt cct cgt 11184  
 Arg Leu Leu Thr Gly Asn Thr Pro Trp Pro Glu Ser Gly Arg Pro Arg  
 3715 3720 3725

cga gcg ggg gtg tcg tcg ttc ggg atc agc ggc acc aac gca cac ctc 11232  
 Arg Ala Gly Val Ser Ser Phe Gly Ile Ser Gly Thr Asn Ala His Leu  
 3730 3735 3740

atc ctc gaa caa cca ccg tcg gaa cca gcg gag atc gac caa tcg gat 11280  
 Ile Leu Glu Gln Pro Pro Ser Glu Pro Ala Glu Ile Asp Gln Ser Asp  
 3745 3750 3755 3760

cgg cgg gtc act gcg cat cca gcg gtg atc ccg tgg atg ttg tcg gct 11328  
 Arg Arg Val Thr Ala His Pro Ala Val Ile Pro Trp Met Leu Ser Ala  
 3765 3770 3775

agg agt ctc gca gcg ctg cag gcc caa gcg gct gcg ctg cag gcc cgg 11376  
 Arg Ser Leu Ala Ala Leu Gln Ala Gln Ala Ala Ala Leu Gln Ala Arg  
 3780 3785 3790

ctg gac cgg ggt cct ggc gct tct ccg ctg gat ttg ggg tat tca ctc 11424  
 Leu Asp Arg Gly Pro Gly Ala Ser Pro Leu Asp Leu Gly Tyr Ser Leu  
 3795 3800 3805

gcg acc act cgt tct gtg ctg gac gaa cgc gcc gtc gtg tgg ggt gcc 11472  
 Ala Thr Thr Arg Ser Val Leu Asp Glu Arg Ala Val Val Trp Gly Ala



3810

3815

3820

gat cgg gag gca ctg ctg tcc agg ctg gca gcg ctc gcc gat ggc cgg 11520  
 Asp Arg Glu Ala Leu Leu Ser Arg Leu Ala Ala Leu Ala Asp Gly Arg  
 3825 3830 3835 3840

acg gcg ccg ggg gtg ata acg ggc tct gcg aat tcc ggt ggc cgc atc 11568  
 Thr Ala Pro Gly Val Ile Thr Gly Ser Ala Asn Ser Gly Gly Arg Ile  
 3845 3850 3855

gga ttc gtt ttt tcc ggt cag ggc agt cag tgg ctg ggg atg gga aag 11616  
 Gly Phe Val Phe Ser Gly Gln Gly Ser Gln Trp Leu Gly Met Gly Lys  
 3860 3865 3870

gcg ttg tgc gcg gct ttc ccg gcg ttc gcg gac gcc ttc gag gaa gcc 11664  
 Ala Leu Cys Ala Ala Phe Pro Ala Phe Ala Asp Ala Phe Glu Glu Ala  
 3875 3880 3885

tgc gac gcg cta agc gca cac ctg ggc gcg gac gtt cgg ggt gtg ctg 11712  
 Cys Asp Ala Leu Ser Ala His Leu Gly Ala Asp Val Arg Gly Val Leu  
 3890 3895 3900

ttc ggt gct gat gag cag atg ctc gac ccg acg ctg tgg gcg cag tcg 11760  
 Phe Gly Ala Asp Glu Gln Met Leu Asp Arg Thr Leu Trp Ala Gln Ser  
 3905 3910 3915 3920

ggg atc ttc gcg gtt caa gtc ggc ctc ctg gga ttg ctg agg tcg tgg 11808  
 Gly Ile Phe Ala Val Gln Val Gly Leu Leu Gly Leu Leu Arg Ser Trp  
 3925 3930 3935

ggc gtg ccg ccg gcc gcg gtg ctg ggg cac tcg gtc ggc gag ttg gct 11856  
 Gly Val Arg Pro Ala Ala Val Leu Gly His Ser Val Gly Glu Leu Ala  
 3940 3945 3950

gcg gcg cac gcg gct ggt gtg ttg tcc ttg ccg gac gct gca ccg ttg 11904  
 Ala Ala His Ala Ala Gly Val Leu Ser Leu Pro Asp Ala Ala Arg Leu  
 3955 3960 3965

gtt gcg gct ccg gcc cac ctg atg cag gca ttg ccc acc ggc ggc gca 11952  
 Val Ala Ala Arg Ala His Leu Met Gln Ala Leu Pro Thr Gly Gly Ala  
 3970 3975 3980

atg ctc gcg gtc gcc acc agc gag gcg gcg gtc gga ccg ctg ctt tcc 12000  
 Met Leu Ala Val Ala Thr Ser Glu Ala Ala Val Gly Pro Leu Leu Ser  
 3985 3990 3995 4000

ggg gtg tgc gat ccg gtc agc atc gct gcg atc aac ggc ccc gag tcg 12048  
 Gly Val Cys Asp Arg Val Ser Ile Ala Ala Ile Asn Gly Pro Glu Ser

4005

4010

4015

gta gtg ctc tcc ggc gac cgc gat gtg ctc gtg gag ctc gca ggc gaa 12096  
 Val Val Leu Ser Gly Asp Arg Asp Val Leu Val Glu Leu Ala Gly Glu  
 4020 4025 4030

ttc gat gcc cga ggg ctt agg acc aaa tgg ttg cgg gtc tcc cat gct 12144  
 Phe Asp Ala Arg Gly Leu Arg Thr Lys Trp Leu Arg Val Ser His Ala  
 4035 4040 4045

ttc cac tcg cac cgg atg gaa ccg att ctg gac gag tac gcg gaa acc 12192  
 Phe His Ser His Arg Met Glu Pro Ile Leu Asp Glu Tyr Ala Glu Thr  
 4050 4055 4060

gcc agg tgc gtc gag ttc ggt gaa ccg gtg gtg ccg atc gtc tcc gcc 12240  
 Ala Arg Cys Val Glu Phe Gly Glu Pro Val Val Pro Ile Val Ser Ala  
 4065 4070 4075 4080

gcg acc ggt gcg ctg gac acc acc gga ctg atg tgc gcg gcc gac tac 12288  
 Ala Thr Gly Ala Leu Asp Thr Thr Gly Leu Met Cys Ala Ala Asp Tyr  
 4085 4090 4095

tgg acg cgc caa gtg cgt gat cct gtc cgc ttc gga gac ggt gtc cgg 12336  
 Trp Thr Arg Gln Val Arg Asp Pro Val Arg Phe Gly Asp Gly Val Arg  
 4100 4105 4110

gcg ctc gtc ggc caa ggc gtg gac acg atc gtc gag ttc ggc ccg gac 12384  
 Ala Leu Val Gly Gln Gly Val Asp Thr Ile Val Glu Phe Gly Pro Asp  
 4115 4120 4125

ggg gcg ttg tcg gcc ctg gtc gag cag tgc ttg gcc ggg tcc gac cag 12432  
 Gly Ala Leu Ser Ala Leu Val Glu Gln Cys Leu Ala Gly Ser Asp Gln  
 4130 4135 4140

gct ggg agg gtg gcg gcg atc ccg ctg atg cgc agg gac cgc gat gag 12480  
 Ala Gly Arg Val Ala Ala Ile Pro Leu Met Arg Arg Asp Arg Asp Glu  
 4145 4150 4155 4160

gtc gag acc gcg gtg gcg gcc ctg gcg cac gtg cac gtc cgc ggt ggt 12528  
 Val Glu Thr Ala Val Ala Ala Leu Ala His Val His Val Arg Gly Gly  
 4165 4170 4175

gcg gtg gac tgg tcg gct tgc ttc gcc ggc acc ggc gcc cgc acc gtc 12576  
 Ala Val Asp Trp Ser Ala Cys Phe Ala Gly Thr Gly Ala Arg Thr Val  
 4180 4185 4190

gag ttg ccc acc tac gcc ttc caa cgc cag ccg tac tgg ctg gcc ggg 12624  
 Glu Leu Pro Thr Tyr Ala Phe Gln Arg Gln Arg Tyr Trp Leu Ala Gly

4195	4200	4205	
caa gcg gac ggg cgc ggc ggc gat gtg gtt gcc gac ccg gtc gac gcg Gln Ala Asp Gly Arg Gly Gly Asp Val Val Ala Asp Pro Val Asp Ala 4210 4215 4220			12672
cgc ttc tgg gag ttg gtc gag cgc gcc gat ccg gaa ccg ttg gtg gat Arg Phe Trp Glu Leu Val Glu Arg Ala Asp Pro Glu Pro Leu Val Asp 4225 4230 4235 4240			12720
gaa ctc tgc atc gac cgg gac cag ccc ttc cgg gag gtg ctg ccc gtt Glu Leu Cys Ile Asp Arg Asp Gln Pro Phe Arg Glu Val Leu Pro Val 4245 4250 4255			12768
ctg gct tcc tgg cgc gag aaa caa cgc cag gag gcc ctc gcg gat tcc Leu Ala Ser Trp Arg Glu Lys Gln Arg Gln Glu Ala Leu Ala Asp Ser 4260 4265 4270			12816
tgg cgc tac cag gtg cgc tgg agg tcc gtc gag gtg ccg tcc gca gcc Trp Arg Tyr Gln Val Arg Trp Arg Ser Val Glu Val Pro Ser Ala Ala 4275 4280 4285			12864
gcc ctc cgg ggc gtg tgg ctg gtg gtg ctt cca gct gac gtg ccc cga Ala Leu Arg Gly Val Trp Leu Val Val Leu Pro Ala Asp Val Pro Arg 4290 4295 4300			12912
gat caa ccg gcg gtc gtc atc gac gcg ctg atc gcg cgc ggc gcc gag Asp Gln Pro Ala Val Val Ile Asp Ala Leu Ile Ala Arg Gly Ala Glu 4305 4310 4315 4320			12960
gtc gcg gtc ctg gaa ttg acc gag cag gac ctc caa cgc agt gcg ctt Val Ala Val Leu Glu Leu Thr Glu Gln Asp Leu Gln Arg Ser Ala Leu 4325 4330 4335			13008
gtg gac aag gtg cgc gcc gtc att gcg gac cgc acc gag gtg acg ggt Val Asp Lys Val Arg Ala Val Ile Ala Asp Arg Thr Glu Val Thr Gly 4340 4345 4350			13056
gtg ttg tct ctg ttg gcg atg gac ggc atg ccc tgc gcg gcg cat ccg Val Leu Ser Leu Leu Ala Met Asp Gly Met Pro Cys Ala Ala His Pro 4355 4360 4365			13104
cac ctg tcc cgt ggt gtc gcc gct acc gtg atc ctg acg cag gtg ttg His Leu Ser Arg Gly Val Ala Ala Thr Val Ile Leu Thr Gln Val Leu 4370 4375 4380			13152
ggc gat gcg ggt gtt tcc gcc ccg ctg tgg ctg gcc acg acc ggt ggc Gly Asp Ala Gly Val Ser Ala Pro Leu Trp Leu Ala Thr Thr Gly Gly			13200

4385	4390	4395	4400
gtc gag gcc ggg acc gag gac ggt ccg gcc gat ccg gac cac ggc ttg			13248
Val Glu Ala Gly Thr Glu Asp Gly Pro Ala Asp Pro Asp His Gly Leu			
4405	4410	4415	
atc tgg ggg ctc ggc agg gtc gtc ggc ctt gaa cat ccg cag tgg tgg			13296
Ile Trp Gly Leu Gly Arg Val Val Gly Leu Glu His Pro Gln Trp Trp			
4420	4425	4430	
ggt ggc ctg atc gac ctt ccg gag aca ctg gac gag acg tcc cgg aac			13344
Gly Gly Leu Ile Asp Leu Pro Glu Thr Leu Asp Glu Thr Ser Arg Asn			
4435	4440	4445	
ggg ttg gtg gcc gca ctc gcc ggg acg gcg gcc gaa gat cag ctc gcc			13392
Gly Leu Val Ala Ala Leu Ala Gly Thr Ala Ala Glu Asp Gln Leu Ala			
4450	4455	4460	
gtg cgt tca tcc ggg ttg ttc gtt cgc aga gtg gtg cgc gca gcg cgg			13440
Val Arg Ser Ser Gly Leu Phe Val Arg Arg Val Val Arg Ala Ala Arg			
4465	4470	4475	4480
aac ccc cgg tca gag aca tgg cgt agc cgg gga acg gtc ctc atc acg			13488
Asn Pro Arg Ser Glu Thr Trp Arg Ser Arg Gly Thr Val Leu Ile Thr			
4485	4490	4495	
ggc gga aca ggc gcg ctc ggt gcc gag gtc gca cga tgg ctg gcc cgg			13536
Gly Gly Thr Gly Ala Leu Gly Ala Glu Val Ala Arg Trp Leu Ala Arg			
4500	4505	4510	
cgg gga gct gag cac ctg gtg ttg atc agt cgc cgc ggc ccg gaa gct			13584
Arg Gly Ala Glu His Leu Val Leu Ile Ser Arg Arg Gly Pro Glu Ala			
4515	4520	4525	
ccc ggc gca gcg gac cta ggg gcc gag ctg act gaa ctc ggc gtg aaa			13632
Pro Gly Ala Ala Asp Leu Gly Ala Glu Leu Thr Glu Leu Gly Val Lys			
4530	4535	4540	
gtc aca gtc ttg gcc tgc gat gtg acg gac cgc gac gag ctg gcg gcg			13680
Val Thr Val Leu Ala Cys Asp Val Thr Asp Arg Asp Glu Leu Ala Ala			
4545	4550	4555	4560
gtg ctg gcg gcc gtt ccc acg gag tat ccg ctg tcg gcg gtc gtg cac			13728
Val Leu Ala Ala Val Pro Thr Glu Tyr Pro Leu Ser Ala Val Val His			
4565	4570	4575	
acc gcc ggc gtc ggg acg cct gcg aac ctg gcc gag acg acc ttg gcg			13776
Thr Ala Gly Val Gly Thr Pro Ala Asn Leu Ala Glu Thr Thr Leu Ala			

4580	4585	4590	
cag ttc gcc gac gtg ttg tcg gcc aag gtc gtc ggc gcg gcg aac ctg			13824
Gln Phe Ala Asp Val Leu Ser Ala Lys Val Val Gly Ala Ala Asn Leu			
4595	4600	4605	
gac cgg ctg ctt ggc ggg caa ccg ttg gac gcc ttc gtg ctg ttc tcc			13872
Asp Arg Leu Leu Gly Gly Gln Pro Leu Asp Ala Phe Val Leu Phe Ser			
4610	4615	4620	
tcg atc tcg gga gtt tgg gga gcc ggc ggc caa gga gcc tat tcg gcc			13920
Ser Ile Ser Gly Val Trp Gly Ala Gly Gly Gln Gly Ala Tyr Ser Ala			
4625	4630	4635	4640
gcc aat gcg tat ctc gat gcc ctt gcc gag cgc cga cgg gct tgc ggg			13968
Ala Asn Ala Tyr Leu Asp Ala Leu Ala Glu Arg Arg Arg Ala Cys Gly			
4645	4650	4655	
cgg ccg gcg acg tgc atc gcc tgg ggt ccg tgg gcg ggt gcg ggc atg			14016
Arg Pro Ala Thr Cys Ile Ala Trp Gly Pro Trp Ala Gly Ala Gly Met			
4660	4665	4670	
gcc gtt cag gaa ggt aac gag gcg cat ctc cgc cga agg ggc ctg gta			14064
Ala Val Gln Glu Gly Asn Glu Ala His Leu Arg Arg Arg Gly Leu Val			
4675	4680	4685	
ccg atg gaa ccg cag tcg gcc ctc ttc gcg ctg caa cag gcc ctg tcc			14112
Pro Met Glu Pro Gln Ser Ala Leu Phe Ala Leu Gln Gln Ala Leu Ser			
4690	4695	4700	
caa cga gaa acc gcc atc acc gtc gca gat gtg gac tgg gag cga ttc			14160
Gln Arg Glu Thr Ala Ile Thr Val Ala Asp Val Asp Trp Glu Arg Phe			
4705	4710	4715	4720
gcc gcc tct ttc acc gcg gcc cgc ccg cga cca ctg ttg gaa gag atc			14208
Ala Ala Ser Phe Thr Ala Ala Arg Pro Arg Pro Leu Leu Glu Glu Ile			
4725	4730	4735	
gtg gat cta cgg ccc gac acc gag acc gag gag aag cac ggt gcc ggc			14256
Val Asp Leu Arg Pro Asp Thr Glu Thr Glu Glu Lys His Gly Ala Gly			
4740	4745	4750	
gag ctg ggg cag cag ctg gcc gca ctg ccg ccc gct gag cgc gga cac			14304
Glu Leu Gly Gln Gln Leu Ala Ala Leu Pro Pro Ala Glu Arg Gly His			
4755	4760	4765	
ctg ctg ctg gag gtg gtg ctg gcg gaa acc gcc agc acc ctg ggg cac			14352
Leu Leu Leu Glu Val Val Leu Ala Glu Thr Ala Ser Thr Leu Gly His			

4770

4775

4780

gat tcg gcg gag gct gtg caa ccc gat cgg acc ttc gcc gaa ctg ggc 14400  
 Asp Ser Ala Glu Ala Val Gln Pro Asp Arg Thr Phe Ala Glu Leu Gly  
 4785 4790 4795 4800

ttc gat tcg ctg acc gcg gta gag ctg cgc aac agg ttg aac gcg gtg 14448  
 Phe Asp Ser Leu Thr Ala Val Glu Leu Arg Asn Arg Leu Asn Ala Val  
 4805 4810 4815

acc ggg ctt cgc ctg ccg ccg acg ctg gtt ttc gac cac ccg acg ccg 14496  
 Thr Gly Leu Arg Leu Pro Pro Thr Leu Val Phe Asp His Pro Thr Pro  
 4820 4825 4830

ctg gcg ttg tcc gaa cag ttg gtt ccg gcc ctg gtc gcg gag ccg gac 14544  
 Leu Ala Leu Ser Glu Gln Leu Val Pro Ala Leu Val Ala Glu Pro Asp  
 4835 4840 4845

aac ggc atc gaa tcg ctg ctc gcc gag ctc gac agg ctg gat acc acg 14592  
 Asn Gly Ile Glu Ser Leu Leu Ala Glu Leu Asp Arg Leu Asp Thr Thr  
 4850 4855 4860

ttg gcg caa ggg cct tcg atc cca ctg gaa gac cag gcc aag gtg gcg 14640  
 Leu Ala Gln Gly Pro Ser Ile Pro Leu Glu Asp Gln Ala Lys Val Ala  
 4865 4870 4875 4880

gag cgc ttg cac gca ctc ctc gcc aag tgg gac ggg gcg cgt gac ggc 14688  
 Glu Arg Leu His Ala Leu Leu Ala Lys Trp Asp Gly Ala Arg Asp Gly  
 4885 4890 4895

acg gcc aga gcg acg tca ccc caa tcg ctg acg gcg gcc acg gac gac 14736  
 Thr Ala Arg Ala Thr Ser Pro Gln Ser Leu Thr Ala Ala Thr Asp Asp  
 4900 4905 4910

gaa atc ttc gac ctc atc gac cgg aag ttc cgg cgc tga 14775  
 Glu Ile Phe Asp Leu Ile Asp Arg Lys Phe Arg Arg  
 4915 4920

&lt;210&gt; 48

&lt;211&gt; 4924

&lt;212&gt; PRT

&lt;213&gt; Saccharopolyspora spinosa

&lt;400&gt; 48

Met Ala Asn Glu Glu Lys Leu Phe Gly Tyr Leu Lys Lys Val Thr Ala  
 1 5 10 15

Asp Leu His Gln Thr Arg Gln Arg Leu Leu Ala Ala Glu Ser Arg Ser  
20 25 30

Gln Glu Pro Ile Ala Ile Val Ser Ala Ser Cys Arg Leu Pro Gly Gly  
35 40 45

Val Asp Ser Pro Glu Ala Leu Trp Gln Leu Val Arg Thr Gly Thr Asp  
50 55 60

Ala Ile Ser Glu Phe Pro Ala Asp Arg Gly Trp Asp Leu Gly Arg Leu  
65 70 75 80

Tyr Asp Pro Asp Pro Asn His Gln Gly Thr Ser Tyr Thr Arg Ala Gly  
85 90 95

Gly Phe Leu Ala Gly Ala Gly Asp Phe Asp Pro Ala Met Phe Gly Ile  
100 105 110

Ser Pro Arg Glu Ala Leu Ala Met Asp Pro Gln Gln Arg Leu Leu Leu  
115 120 125

Glu Leu Ser Trp Glu Ala Leu Glu Arg Ala Gly Ile Asp Pro Thr Ser  
130 135 140

Leu Arg Gly Ser Lys Thr Gly Val Phe Gly Gly Val Thr Pro Gln Glu  
145 150 155 160

Tyr Gly Pro Ser Leu Gln Glu Met Ser Arg Asn Ala Gly Gly Phe Gly  
165 170 175

Leu Thr Gly Arg Met Val Ser Val Ala Ser Gly Arg Val Ala Tyr Ser  
180 185 190

Phe Gly Phe Glu Gly Pro Ala Val Thr Val Asp Thr Ala Cys Ser Ser  
195 200 205

Ser Leu Val Ala Leu His Leu Ala Cys Gln Ser Leu Arg Ser Gly Glu  
210 215 220

Cys Asp Leu Ala Leu Ala Gly Gly Val Thr Val Met Ala Thr Pro Ala  
225 230 235 240

Thr Phe Val Glu Phe Ser Arg Gln Arg Gly Leu Ala Pro Asp Gly Arg  
245 250 255

Cys Lys Ser Phe Ala Ala Ala Ala Asp Gly Thr Gly Trp Gly Glu Gly  
260 265 270

Ala Gly Leu Val Leu Leu Glu Arg Leu Ser Asp Ala Arg Arg Asn Gly  
 275 280 285

His Glu Val Leu Ala Val Val Arg Gly Ser Ala Val Asn Gln Asp Gly  
 290 295 300

Ala Ser Asn Gly Leu Thr Ala Pro Asn Gly Pro Ser Gln Gln Arg Val  
 305 310 315 320

Ile Thr Gln Ala Leu Ala Ser Ala Gly Leu Ser Val Ser Asp Val Asp  
 325 330 335

Ala Val Glu Ala His Gly Thr Gly Thr Thr Leu Gly Asp Pro Ile Glu  
 340 345 350

Ala Gln Ala Leu Ile Ala Thr Tyr Gly Gln Gly Arg Glu Lys Asp Arg  
 355 360 365

Pro Leu Trp Leu Gly Ser Val Lys Ser Asn Ile Gly His Thr Gln Ala  
 370 375 380

Ala Ala Gly Val Ala Gly Val Ile Lys Met Val Leu Ala Met Arg His  
 385 390 395 400

Gly Gln Leu Pro Ala Thr Leu His Val Asp Glu Pro Thr Ser Ala Val  
 405 410 415

Asp Trp Ser Ala Gly Ser Val Arg Leu Leu Thr Glu Asn Thr Pro Trp  
 420 425 430

Pro Asp Ser Gly Arg Pro Cys Arg Val Gly Val Ser Ser Phe Gly Ile  
 435 440 445

Ser Gly Thr Asn Ala His Val Ile Leu Glu Gln Ser Pro Val Glu Gln  
 450 455 460

Gly Glu Pro Ala Gly Pro Val Glu Gly Glu Arg Glu Pro Asp Val Ala  
 465 470 475 480

Val Pro Val Val Pro Trp Val Leu Ser Gly Lys Thr Pro Glu Ala Ala  
 485 490 495

Arg Ala Gln Ala Glu Arg Val His Ser His Ile Glu Asp Arg Pro Gly  
 500 505 510

Leu Ser Pro Val Asp Val Ala Tyr Ser Leu Gly Met Thr Arg Ala Ala  
 515 520 525



Leu Asp Glu Arg Ala Val Val Leu Gly Ser Asp Arg Ala Ala Leu Leu  
530 535 540

Thr Gly Leu Arg Ala Phe Ala Asp Gly Cys Asp Ala Pro Glu Val Val  
545 550 555 560

Ser Gly Ser Val Gly Leu Gly Gly Arg Val Gly Phe Val Phe Ser Gly  
565 570 575

Gln Gly Gly Gln Trp Pro Gly Met Gly Arg Gly Leu Tyr Ser Val Phe  
580 585 590

Pro Val Phe Ala Asp Ala Phe Asp Glu Ala Cys Ala Glu Leu Asp Ala  
595 600 605

His Leu Gly Gln Glu Leu Arg Val Arg Asp Val Val Phe Gly Ser Gln  
610 615 620

Ala Trp Leu Leu Asp Arg Thr Val Trp Ala Gln Ser Gly Leu Phe Ala  
625 630 635 640

Leu Gln Ile Gly Leu Leu Arg Leu Leu Gly Ser Trp Gly Val Arg Pro  
645 650 655

Asp Val Val Leu Gly His Ser Val Gly Glu Leu Ala Ala Val His Ala  
660 665 670

Ala Gly Val Leu Ser Leu Ser Glu Ala Ala Arg Leu Val Ala Gly Arg  
675 680 685

Ala Arg Leu Met Gln Ala Leu Pro Ser Gly Gly Ala Met Leu Ala Val  
690 695 700

Ala Thr Gly Glu Phe Gln Val Asp Pro Leu Leu Asp Gly Val Arg Asp  
705 710 715 720

Arg Ile Gly Ile Ala Ala Val Asn Gly Pro Glu Ser Val Val Leu Ser  
725 730 735

Gly Asp Arg Glu Leu Leu Thr Glu Ile Ala Asp Arg Leu His Asp Gln  
740 745 750

Gly Cys Arg Thr Arg Trp Leu Arg Val Ser His Ala Phe His Ser Pro  
755 760 765

His Met Glu Pro Met Leu Glu Glu Phe Ala Gln Ile Ser Arg Gly Arg  
770 775 780

Glu Tyr His Ala Pro Glu Leu Pro Ile Ile Ser Thr Leu Ile Gly Glu  
 785 790 795 800

Leu Asp Gly Gly Arg Val Met Gly Thr Pro Glu Tyr Trp Val Arg Gln  
 805 810 815

Val Arg Glu Pro Val Arg Phe Ala Glu Gly Val Gln Ala Leu Val Gly  
 820 825 830

Gln Gly Val Gly Thr Ile Val Glu Leu Gly Pro Asp Gly Ala Leu Ser  
 835 840 845

Thr Leu Val Glu Glu Cys Val Ala Glu Ser Gly Arg Val Ala Gly Ile  
 850 855 860

Pro Leu Met Arg Lys Asp Arg Asp Glu Ala Arg Thr Val Leu Ala Ala  
 865 870 875 880

Leu Ala Gln Ile His Thr Arg Gly Gly Glu Val Asp Trp Arg Ser Phe  
 885 890 895

Phe Ala Gly Thr Gly Ala Lys Gln Val Asp Leu Pro Thr Tyr Ala Phe  
 900 905 910

Gln Arg Gln Arg Tyr Trp Leu Ala Ser Thr Gly Arg Ala Gly Asp Val  
 915 920 925

Thr Ala Ala Gly Leu Ala Glu Ala Asp His Pro Leu Leu Gly Ala Val  
 930 935 940

Val Ala Leu Ala Asp Gly Glu Gly Val Val Leu Thr Gly Arg Leu Thr  
 945 950 955 960

Ala Gly Ser His Pro Trp Leu Ser Asp His Arg Val Leu Gly Glu Ile  
 965 970 975

Val Val Pro Gly Thr Ala Ile Val Glu Leu Val Trp His Val Gly Glu  
 980 985 990

Arg Leu Gly Cys Gly Arg Val Glu Glu Leu Ala Leu Glu Ala Pro Leu  
 995 1000 1005

Ile Leu Pro Asp His Gly Ala Val Gln Val Gln Val Leu Val Gly Pro  
 1010 1015 1020

Pro Gly Glu Ser Gly Ala Arg Ser Val Ala Leu Tyr Ser Cys Pro Gly  
 1025 1030 1035 1040

Glu Ala Ile Glu Pro Glu Trp Lys Lys His Ala Thr Gly Val Leu Leu  
1045 1050 1055

Pro Pro Val Ala Ala Glu Asn His Glu Leu Thr Ala Trp Pro Pro Glu  
1060 1065 1070

Asn Ala Thr Glu Ile Asp Ala Asp Gly Val Tyr Ala Phe Leu Glu Gly  
1075 1080 1085

His Gly Phe Ala Tyr Gly Pro Ala Phe Arg Cys Leu Arg Gly Ala Trp  
1090 1095 1100

Arg Arg Gly Gly Glu Val Phe Ala Glu Val Ala Leu Pro Asp Asp Met  
1105 1110 1115 1120

Gln Ala Gly Val Asp Arg Phe Gly Val His Pro Ala Leu Leu Asp Ala  
1125 1130 1135

Val Leu His Ala Ala Ala Ala Glu Thr Ser Val Val Gln Ser Glu Ala  
1140 1145 1150

Arg Val Pro Phe Ser Trp Arg Gly Val Glu Leu Arg Ala Thr Glu Ser  
1155 1160 1165

Ala Val Val Arg Ala Arg Leu Ser Leu Thr Ser Asp Asp Glu Leu Ser  
1170 1175 1180

Leu Val Ala Val Asp Pro Ala Gly Arg Phe Val Ala Thr Val Asp Ser  
1185 1190 1195 1200

Leu Val Thr Arg Pro Ile Ser Arg Gln Gln Val Arg Ser Gly Ala Ile  
1205 1210 1215

Gly Asp Cys Leu Phe Glu Val Glu Trp His Arg Lys Ala Leu Leu Gly  
1220 1225 1230

Thr Thr Ala Gly Asp Asp Leu Ala Ile Val Gly Asp Gly Pro Ser Trp  
1235 1240 1245

Pro Glu Ser Val Arg Ala Thr Ala Arg Phe Ala Thr Leu Asp Glu Phe  
1250 1255 1260

Arg Ala Ala Val Asp Ser Asp Val Pro Ala Pro Gly Ser Val Leu Val  
1265 1270 1275 1280

Ala Ala Met Ser Ala Glu Glu Val Glu Gly Gly Ser Leu Pro Ser Arg  
1285 1290 1295

Ala Gln Glu Ser Thr Ser Asp Leu Leu Ala Leu Val Gln Ser Trp Leu  
 1300 1305 1310

Ala Asp Glu Arg Phe Ala Glu Ser Gln Leu Val Val Val Thr Arg Ala  
 1315 1320 1325

Ala Val Ser Ala Asp Ser Asp Ser Asp Val Ala Asp Leu Val Gly Ala  
 1330 1335 1340

Ser Ser Trp Gly Leu Leu Ser Ser Ala Gln Ser Glu Asn Pro Gly Arg  
 1345 1350 1355 1360

Phe Val Leu Val Asp Val Asp Gly Thr Pro Glu Ser Trp Gln Ala Leu  
 1365 1370 1375

Pro Ala Ala Val Arg Ala Gly Glu Pro Gln Leu Ala Leu Arg Arg Gly  
 1380 1385 1390

Val Ala Leu Val Pro Arg Leu Ala Arg Leu Thr Val Arg Glu Glu Gly  
 1395 1400 1405

Ser Ser Pro Gln Leu Asp Thr Asp Gly Thr Val Leu Ile Thr Gly Gly  
 1410 1415 1420

Thr Gly Ala Leu Gly Gly Val Val Ala Arg His Leu Val Glu Glu His  
 1425 1430 1435 1440

Gly Ile Arg Arg Leu Val Leu Ala Gly Arg Arg Gly Trp Asn Ala Pro  
 1445 1450 1455

Gly Val His Glu Leu Val Asp Glu Leu Ala Arg Ala Gly Ala Val Val  
 1460 1465 1470

Val Val Ala Cys Asp Val Ala Asp Arg Thr Asp Leu Glu His Val  
 1475 1480 1485

Leu Ala Ala Ile Pro Val Asp Trp Pro Leu Arg Gly Ile Val His Thr  
 1490 1495 1500

Ala Gly Val Leu Ala Asp Gly Val Ile Gly Ser Leu Ser Ala Ala Asp  
 1505 1510 1515 1520

Val Gly Thr Val Phe Ala Pro Lys Val Thr Gly Ala Trp His Leu His  
 1525 1530 1535

Glu Leu Thr Arg Asp Leu Asp Leu Ser Phe Phe Val Leu Phe Ser Ser  
 1540 1545 1550

Phe Ser Gly Ile Ala Gly Ala Ala Gly Gln Ala Asn Tyr Ala Ala Ala  
1555 1560 1565

Asn Thr Phe Leu Asp Ala Leu Ala Arg Tyr Arg Arg Ala Arg Gly Leu  
1570 1575 1580

Pro Gly Leu Ser Leu Ala Trp Gly Leu Trp Ala Gln Pro Ser Gly Met  
1585 1590 1595 1600

Thr Ser Gly Leu Asp Ala Ala Ser Val Glu Arg Leu Ala Arg Thr Gly  
1605 1610 1615

Ile Ala Glu Leu Ser Thr Glu Asp Gly Leu Arg Leu Phe Asp Ala Ala  
1620 1625 1630

Phe Ala Lys Asp Arg Ala Cys Val Val Ala Ala Arg Leu Asp Arg Ala  
1635 1640 1645

Leu Leu Val Gly Asn Gly Arg Ser His Ala Ile Pro Ala Leu Leu Ser  
1650 1655 1660

Ala Leu Val Pro Val Arg Gly Gly Val Ala Arg Lys Thr Ala Asn Ser  
1665 1670 1675 1680

Gln Ala Ala Asp Glu Asp Ala Leu Leu Gly Leu Val Arg Glu His Val  
1685 1690 1695

Ser Ala Val Leu Gly Tyr Ser Gly Ala Val Glu Val Gly Gly Asp Arg  
1700 1705 1710

Ala Phe Arg Asp Leu Gly Phe Asp Ser Leu Ser Gly Val Glu Leu Arg  
1715 1720 1725

Arg Leu Ala Gly Val Leu Gly Val Arg Leu Pro Ala Thr Ala Val  
1730 1735 1740

Phe Asp Tyr Pro Thr Pro Arg Ala Leu Ala Arg Phe Leu His Gln Glu  
1745 1750 1755 1760

Leu Ala Gly Glu Val Ala Ser Thr Ser Thr Pro Val Thr Arg Ala Ala  
1765 1770 1775

Ser Ala Glu Glu Asp Leu Val Ala Ile Val Gly Met Gly Cys Arg Phe  
1780 1785 1790

Pro Gly Gly Val Ser Ser Pro Glu Glu Leu Trp Arg Leu Val Ala Gly  
1795 1800 1805

Gly Val Asp Ala Val Ala Gly Phe Pro Asp Asp Arg Gly Trp Asp Leu  
1810 1815 1820

Ala Ala Leu Tyr Asp Pro Asp Pro Asp Arg Leu Gly Thr Ser Tyr Val  
825 1830 1835 1840

Cys Glu Gly Gly Phe Leu Arg Asp Ala Ala Glu Phe Asp Ala Asp Met  
1845 1850 1855

Phe Gly Ile Ser Pro Arg Glu Ala Leu Ala Met Asp Pro Gln Gln Arg  
1860 1865 1870

Leu Leu Leu Glu Val Ala Trp Glu Thr Leu Glu Arg Ala Gly Ile Asp  
1875 1880 1885

Pro Phe Ser Leu His Gly Ser Arg Thr Gly Val Phe Ala Gly Leu Met  
890 1895 1900

His Asp Tyr Gly Ala Arg Phe Ile Thr Arg Ala Pro Glu Gly Phe  
905 1910 1915 1920

Glu Gly His Leu Gly Thr Gly Asn Ala Gly Ser Val Leu Ser Gly Arg  
1925 1930 1935

Val Ala Tyr Ser Phe Gly Phe Glu Gly Pro Ala Val Thr Val Asp Thr  
1940 1945 1950

Ala Cys Ser Ser Ser Leu Val Ala Leu His Leu Ala Gly Gln Ala Leu  
1955 1960 1965

Arg Ala Gly Glu Cys Glu Phe Ala Leu Ala Gly Gly Val Thr Val Met  
1970 1975 1980

Thr Pro Thr Thr Phe Val Glu Phe Ser Arg Gln Arg Gly Leu Ala  
1985 1990 1995 2000

Pro Asp Gly Arg Cys Lys Ser Phe Ala Ala Ala Ala Asp Gly Thr Gly  
2005 2010 2015

Trp Gly Glu Gly Ala Gly Leu Val Leu Leu Glu Arg Leu Ser Asp Ala  
2020 2025 2030

Arg Arg Asn Gly His Glu Val Leu Ala Val Val Arg Gly Ser Ala Val  
2035 2040 2045

Asn Gln Asp Gly Ala Ser Asn Gly Leu Thr Ala Pro Asn Gly Pro Ser  
2050 2055 2060

Gln Gln Arg Val Ile Thr Gln Ala Leu Thr Ser Ala Gly Leu Ser Val  
065 2070 2075 2080

Ser Asp Val Asp Ala Val Glu Ala His Gly Thr Gly Thr Arg Leu Gly  
2085 2090 2095

Asp Pro Ile Glu Ala Gln Ala Leu Ile Ala Thr Tyr Gly Arg Asp Arg  
2100 2105 2110

Asp Pro Gly Arg Pro Leu Trp Leu Gly Ser Val Lys Ser Asn Ile Gly  
2115 2120 2125

His Thr Gln Ala Ala Ala Gly Val Ala Gly Val Ile Lys Met Val Met  
2130 2135 2140

Ala Met Arg Gln Gly Glu Leu Pro Arg Thr Leu His Val Asp Glu Pro  
145 2150 2155 2160

Ser Ala Gln Val Asp Trp Ser Ala Gly Thr Val Gln Leu Leu Thr Glu  
2165 2170 2175

Asn Thr Pro Trp Pro Asp Ser Gly Arg Leu Arg Arg Ala Gly Val Ser  
2180 2185 2190

Ser Phe Gly Ile Ser Gly Thr Asn Ala His Leu Ile Leu Glu Gln Pro  
2195 2200 2205

Pro Arg Glu Ser Gln Arg Ser Thr Glu Pro Asp Ser Gly Ser Val Arg  
2210 2215 2220

Asp Phe Pro Val Val Pro Trp Met Val Ser Gly Lys Thr Pro Glu Ala  
225 2230 2235 2240

Leu Ser Ala Gln Ala Asp Ala Leu Met Ser Tyr Leu Ser Asn Arg Val  
2245 2250 2255

Asp Ala Ser Pro Arg Asp Ile Gly Tyr Ser Leu Ala Val Thr Arg Pro  
2260 2265 2270

Ala Leu Asp His Arg Ala Val Val Leu Gly Ala Asp Arg Ala Ala Leu  
2275 2280 2285

Leu Pro Gly Leu Lys Ala Leu Ala Val Ser Asn Asp Ala Ala Glu Val  
2290 2295 2300

Ile Thr Gly Thr Arg Ala Ala Gly Pro Val Gly Phe Val Phe Ser Gly  
305 2310 2315 2320

Gln Gly Gly Gln Thr Pro Gly Met Gly Ser Gly Leu His Ser Ala Phe  
2325 2330 2335

Pro Val Phe Ala Asp Ala Phe Asp Glu Ala Cys Cys Glu Leu Asp Ala  
2340 2345 2350

His Leu Gly Gln Met Ala Arg Leu Arg Asp Val Leu Ser Gly Ser Asp  
2355 2360 2365

Thr Gln Leu Leu Asp Gln Thr Leu Trp Ala Gln Pro Gly Leu Phe Ala  
2370 2375 2380

Leu Gln Val Gly Leu Trp Glu Leu Leu Gly Ser Trp Gly Val Arg Pro  
385 2390 2395 2400

Ala Val Val Leu Gly His Ser Val Gly Glu Leu Ala Ala Ala Phe Ala  
2405 2410 2415

Ala Gly Val Leu Ser Leu Arg Asp Ala Ala Arg Leu Val Ala Gly Arg  
2420 2425 2430

Ala Arg Leu Met Gln Ala Leu Pro Thr Gly Gly Ala Met Leu Ala Ala  
2435 2440 2445

Ala Ala Gly Glu Glu Gln Leu Arg Pro Leu Leu Ala Asp Cys Gly Asp  
2450 2455 2460

Arg Val Gly Ile Ala Ala Val Asn Ala Pro Gly Ser Val Val Leu Ser  
465 2470 2475 2480

Gly Asp Arg Asp Val Leu Asp Asp Ile Ala Gly Arg Leu Asp Gly Gln  
2485 2490 2495

Gly Ile Arg Ser Arg Trp Leu Arg Val Ser His Ala Phe His Ser His  
2500 2505 2510

Arg Met Asp Pro Met Leu Ala Glu Phe Thr Glu Ile Ala Arg Ser Val  
2515 2520 2525

Asp Tyr Arg Ser Ser Gly Leu Pro Ile Val Ser Thr Leu Thr Gly Glu  
2530 2535 2540

Leu Asp Glu Val Gly Met Pro Ala Thr Pro Glu Tyr Trp Val Arg Gln  
545 2550 2555 2560

Val Arg Glu Pro Val Arg Phe Ala Asp Gly Val Ala Ala Leu Ala Ala  
2565 2570 2575



His Gly Val Ser Thr Val Val Glu Val Gly Pro Asp Gly Val Leu Ser  
2580 2585 2590

Ala Leu Val Gln Glu Cys Ala Ala Gly Ser Asp Gln Gly Gly Arg Val  
2595 2600 2605

Ala Ala Val Pro Leu Met Arg Ser Asn Arg Asp Glu Ala His Thr Val  
2610 2615 2620

Thr Thr Ala Leu Ala Gln Ile His Val Arg Gly Ala Glu Val Asp Trp  
625 2630 2635 2640

Arg Ser Phe Phe Ala Gly Thr Gly Ala Lys Gln Val Glu Leu Pro Thr  
2645 2650 2655

Tyr Ala Phe Gln Arg Gln Arg Tyr Trp Leu Asp Ser Pro Ser Glu Pro  
2660 2665 2670

Val Gly Gln Ser Ala Asp Pro Ala Arg Gln Ser Gly Phe Trp Glu Leu  
2675 2680 2685

Val Glu Gln Glu Asp Val Ser Ala Leu Ser Ala Ala Leu His Ile Thr  
2690 2695 2700

Gly Asp His Asp Val Gln Ala Ser Leu Glu Ser Val Val Pro Val Leu  
705 2710 2715 2720

Ser Ser Trp His Arg Arg Ile Arg Asn Glu Ser Leu Val His Gln Trp  
2725 2730 2735

Arg Tyr Arg Ile Ser Trp His Glu Arg Ala Asp Leu Pro Asp Pro Ser  
2740 2745 2750

Leu Ser Gly Thr Trp Leu Val Val Val Pro Glu Gly Trp Ser Ala Ser  
2755 2760 2765

Arg Gln Val Leu Arg Phe Asn Glu Met Phe Glu Glu Arg Gly Cys Pro  
2770 2775 2780

Ala Val Leu Phe Glu Leu Ala Gly His Asp Glu Glu Ala Leu Ala Gln  
785 2790 2795 2800

Arg Phe Arg Ser Leu Pro Val Ala Ser Gly Gly Ile Ser Gly Val Leu  
2805 2810 2815

Ser Leu Leu Ala Leu Asp Glu Ser Pro Ser Ser Pro Asn Ala Ala Leu  
2820 2825 2830

Pro Asn Gly Ala I      sn Ser Leu Val Leu Leu Arg Ala Leu Arg Ala  
 2835                      2840                      2845

Ala Asp Val Ser Ala Pro Leu Trp Leu Ala Thr Cys Gly Gly Val Ala  
 2850                      2855                      2860

Val Gly Asp Val Pro Val Asn Pro Gly Gln Ala Leu Val Trp Gly Leu  
 865                      2870                      2875                      2880

Gly Arg Val Val Gly Leu Glu His Pro Ala Trp Trp Gly Gly Leu Val  
 2885                      2890                      2895

Asp Val Pro Cys Leu Leu Asp Glu Asp Ala Arg Glu Arg Leu Ser Val  
 2900                      2905                      2910

Val Leu Ala Gly Leu Gly Glu Asp Glu Ile Ala Val Arg Pro Gly Gly  
 2915                      2920                      2925

Val Phe Val Arg Arg Leu Glu Arg Ala Gly Ala Ala Ser Gly Ala Gly  
 2930                      2935                      2940

Ser Val Trp Arg Pro Arg Gly Thr Val Leu Val Thr Gly Gly Thr Gly  
 945                      2950                      2955                      2960

Gly Leu Gly Ala His Val Ala Arg Trp Leu Ala Gly Ala Gly Ala Glu  
 2965                      2970                      2975

His Val Val Leu Thr Ser Arg Arg Gly Ala Ala Ala Pro Gly Ala Gly  
 2980                      2985                      2990

Asp Leu Arg Ala Glu Leu Glu Ala Leu Gly Ala Arg Val Ser Ile Thr  
 2995                      3000                      3005

Ala Cys Asp Val Ala Asp Arg Asp Ala Leu Ala Glu Val Leu Ala Thr  
 3010                      3015                      3020

Ile Pro Asp Asp Cys Pro Leu Thr Ala Val Met His Ala Ala Gly Val  
 025                      3030                      3035                      3040

Val Glu Val Gly Asp Val Ala Ser Met Cys Leu Thr Asp Phe Val Gly  
 3045                      3050                      3055

Val Leu Ser Ala Lys Ala Gly Gly Ala Ala Asn Leu Asp Glu Leu Leu  
 3060                      3065                      3070

Ala Asp Val Glu Leu Asp Ala Phe Val Leu Phe Ser Ser Val Ser Gly  
 3075                      3080                      3085

Val Trp Gly Ala Gly Gly Gln Gly Ala Tyr Ala Ala Ala Asn Ala Tyr  
3090 3095 3100

Leu Asp Ala Leu Ala Gln Gln Arg Arg Ala Arg Gly Leu Val Gly Thr  
105 3110 3115 3120

Ala Val Ala Trp Gly Pro Trp Ala Gly Asp Gly Met Ala Ala Gly Glu  
3125 3130 3135

Gly Gly Ala Gln Leu Arg Arg Ala Gly Leu Val Pro Met Ala Ala Asp  
3140 3145 3150

Arg Ala Leu Leu Ala Leu Gln Gly Ala Leu Asp Arg Asp Glu Thr Ser  
3155 3160 3165

Leu Val Val Ala Asp Met Ala Trp Glu Arg Phe Ala Pro Val Phe Ala  
3170 3175 3180

Met Ser Arg Arg Arg Pro Leu Leu Asp Glu Leu Pro Glu Ala Gln Gln  
185 3190 3195 3200

Ala Leu Ala Asp Ala Glu Asn Thr Thr Asp Ala Ala Asp Ser Ala Val  
3205 3210 3215

Pro Leu Pro Arg Leu Ala Gly Met Ala Ala Ala Glu Arg Arg Arg Ala  
3220 3225 3230

Met Leu Asp Leu Val Leu Ala Glu Ala Ser Ile Val Leu Gly His Asn  
3235 3240 3245

Gly Ser Asp Pro Val Gly Pro Asp Arg Ala Phe Gln Glu Leu Gly Phe  
3250 3255 3260

Asp Ser Leu Met Ala Val Glu Leu Arg Asn Arg Leu Gly Glu Ala Thr  
3265 3270 3275 3280

Gly Leu Ser Leu Pro Ala Thr Leu Ile Phe Asp Tyr Pro Ser Pro Ser  
3285 3290 3295

Ala Leu Ala Glu Gln Leu Val Gly Glu Leu Val Gly Ala Gln Pro Ala  
3300 3305 3310

Thr Thr Val Val Ala Gly Ala Asp Pro Val Asp Asp Pro Val Val Val  
3315 3320 3325

Val Ala Met Gly Cys Arg Tyr Pro Gly Asp Val Cys Ser Pro Glu Glu  
3330 3335 3340

Leu Trp Gln Leu Val Ser Ala Gly Arg Asp Ala Val Ser Thr Phe Pro  
345 3350 3355 3360

Val Asp Arg Gly Trp Asp Cys Asn Thr Leu Phe Asp Pro Asp Pro Asp  
3365 3370 3375

Arg Ala Gly Ser Thr Tyr Val Arg Glu Gly Ala Phe Leu Thr Gly Ala  
3380 3385 3390

Asp Arg Phe Asp Ala Gly Phe Phe Gly Ile Ser Pro Arg Glu Ala Arg  
3395 3400 3405

Ala Met Asp Pro Gln Gln Arg Leu Leu Leu Glu Val Ala Trp Glu Val  
3410 3415 3420

Phe Glu Arg Ala Gly Ile Ala Pro Leu Ser Leu Arg Gly Ser Arg Thr  
3425 3430 3435 3440

Gly Val Phe Ala Gly Thr Asn Gly Gln Asp His Gly Ala Lys Val Ala  
3445 3450 3455

Ala Ala Pro Glu Ala Ala Gly His Leu Leu Thr Gly Asn Ala Ala Ser  
3460 3465 3470

Val Leu Ala Gly Arg Leu Ser Tyr Thr Phe Gly Leu Glu Gly Pro Ala  
3475 3480 3485

Val Ala Val Asp Thr Ala Cys Ser Ser Ser Leu Val Ala Leu His Leu  
3490 3495 3500

Ala Cys Gln Ser Leu Arg Ser Gly Glu Cys Asp Met Ala Leu Ala Gly  
3505 3510 3515 3520

Gly Val Thr Val Met Ser Thr Pro Leu Ala Phe Leu Glu Phe Ser Arg  
3525 3530 3535

Gln Arg Gly Leu Ala Pro Asp Gly Arg Cys Lys Ser Phe Ala Ala Ala  
3540 3545 3550

Ala Asp Gly Thr Gly Trp Gly Glu Gly Ala Gly Leu Val Leu Leu Glu  
3555 3560 3565

Arg Leu Ser Asp Ala Arg Arg Asn Gly His Arg Val Leu Ala Val Val  
3570 3575 3580

Arg Gly Ser Ala Val Asn Gln Asp Gly Ala Ser Asn Gly Leu Thr Ala  
3585 3590 3595 3600

Pro Asn Gly Pro Ser Gln Gln Arg Val Ile Arg Gln Ala Leu Ala Asn  
 3605 3610 3615  
 Ala Gly Leu Ser Ala Ser Asp Val Asp Val Val Glu Ala His Gly Thr  
 3620 3625 3630  
 Gly Thr Gly Leu Gly Asp Pro Ile Glu Ala Gln Ala Leu Ile Ala Thr  
 3635 3640 3645  
 Tyr Gly Gln Glu Arg Asp Pro Glu Arg Ala Leu Trp Leu Gly Ser Ile  
 3650 3655 3660  
 Lys Ser Asn Ile Gly His Thr Gln Ala Ala Ala Gly Val Ala Gly Val  
 665 3670 3675 3680  
 Ile Lys Met Val Gln Ala Met Arg His Gly Glu Leu Pro Ala Thr Leu  
 3685 3690 3695  
 His Val Asp Lys Pro Thr Pro Gln Val Asp Trp Ser Ala Gly Ala Val  
 3700 3705 3710  
 Arg Leu Leu Thr Gly Asn Thr Pro Trp Pro Glu Ser Gly Arg Pro Arg  
 3715 3720 3725  
 Arg Ala Gly Val Ser Ser Phe Gly Ile Ser Gly Thr Asn Ala His Leu  
 3730 3735 3740  
 Ile Leu Glu Gln Pro Pro Ser Glu Pro Ala Glu Ile Asp Gln Ser Asp  
 745 3750 3755 3760  
 Arg Arg Val Thr Ala His Pro Ala Val Ile Pro Trp Met Leu Ser Ala  
 3765 3770 3775  
 Arg Ser Leu Ala Ala Leu Gln Ala Gln Ala Ala Ala Leu Gln Ala Arg  
 3780 3785 3790  
 Leu Asp Arg Gly Pro Gly Ala Ser Pro Leu Asp Leu Gly Tyr Ser Leu  
 3795 3800 3805  
 Ala Thr Thr Arg Ser Val Leu Asp Glu Arg Ala Val Val Trp Gly Ala  
 3810 3815 3820  
 Asp Arg Glu Ala Leu Leu Ser Arg Leu Ala Ala Leu Ala Asp Gly Arg  
 825 3830 3835 3840  
 Thr Ala Pro Gly Val Ile Thr Gly Ser Ala Asn Ser Gly Gly Arg Ile  
 3845 3850 3855

Gly Phe Val Phe S Gly Gln Gly Ser Gln Trp Leu G Met Gly Lys  
3860 3865 3870

Ala Leu Cys Ala Ala Phe Pro Ala Phe Ala Asp Ala Phe Glu Glu Ala  
3875 3880 3885

Cys Asp Ala Leu Ser Ala His Leu Gly Ala Asp Val Arg Gly Val Leu  
3890 3895 3900

Phe Gly Ala Asp Glu Gln Met Leu Asp Arg Thr Leu Trp Ala Gln Ser  
905 3910 3915 3920

Gly Ile Phe Ala Val Gln Val Gly Leu Leu Gly Leu Leu Arg Ser Trp  
3925 3930 3935

Gly Val Arg Pro Ala Ala Val Leu Gly His Ser Val Gly Glu Leu Ala  
3940 3945 3950

Ala Ala His Ala Ala Gly Val Leu Ser Leu Pro Asp Ala Ala Arg Leu  
3955 3960 3965

Val Ala Ala Arg Ala His Leu Met Gln Ala Leu Pro Thr Gly Gly Ala  
3970 3975 3980

Met Leu Ala Val Ala Thr Ser Glu Ala Ala Val Gly Pro Leu Leu Ser  
985 3990 3995 4000

Gly Val Cys Asp Arg Val Ser Ile Ala Ala Ile Asn Gly Pro Glu Ser  
4005 4010 4015

Val Val Leu Ser Gly Asp Arg Asp Val Leu Val Glu Leu Ala Gly Glu  
4020 4025 4030

Phe Asp Ala Arg Gly Leu Arg Thr Lys Trp Leu Arg Val Ser His Ala  
4035 4040 4045

Phe His Ser His Arg Met Glu Pro Ile Leu Asp Glu Tyr Ala Glu Thr  
4050 4055 4060

Ala Arg Cys Val Glu Phe Gly Glu Pro Val Val Pro Ile Val Ser Ala  
065 4070 4075 4080

Ala Thr Gly Ala Leu Asp Thr Thr Gly Leu Met Cys Ala Ala Asp Tyr  
4085 4090 4095

Trp Thr Arg Gln Val Arg Asp Pro Val Arg Phe Gly Asp Gly Val Arg  
4100 4105 4110

Ala Leu Val Gly Gln Gly Val Asp Thr Ile Val Glu Phe Gly Pro Asp  
4115 4120 4125

Gly Ala Leu Ser Ala Leu Val Glu Gln Cys Leu Ala Gly Ser Asp Gln  
4130 4135 4140

Ala Gly Arg Val Ala Ala Ile Pro Leu Met Arg Arg Asp Arg Asp Glu  
145 4150 4155 4160

Val Glu Thr Ala Val Ala Ala Leu Ala His Val His Val Arg Gly Gly  
4165 4170 4175

Ala Val Asp Trp Ser Ala Cys Phe Ala Gly Thr Gly Ala Arg Thr Val  
4180 4185 4190

Glu Leu Pro Thr Tyr Ala Phe Gln Arg Gln Arg Tyr Trp Leu Ala Gly  
4195 4200 4205

Gln Ala Asp Gly Arg Gly Gly Asp Val Val Ala Asp Pro Val Asp Ala  
4210 4215 4220

Arg Phe Trp Glu Leu Val Glu Arg Ala Asp Pro Glu Pro Leu Val Asp  
225 4230 4235 4240

Glu Leu Cys Ile Asp Arg Asp Gln Pro Phe Arg Glu Val Leu Pro Val  
4245 4250 4255

Leu Ala Ser Trp Arg Glu Lys Gln Arg Gln Glu Ala Leu Ala Asp Ser  
4260 4265 4270

Trp Arg Tyr Gln Val Arg Trp Arg Ser Val Glu Val Pro Ser Ala Ala  
4275 4280 4285

Ala Leu Arg Gly Val Trp Leu Val Val Leu Pro Ala Asp Val Pro Arg  
4290 4295 4300

Asp Gln Pro Ala Val Val Ile Asp Ala Leu Ile Ala Arg Gly Ala Glu  
305 4310 4315 4320

Val Ala Val Leu Glu Leu Thr Glu Gln Asp Leu Gln Arg Ser Ala Leu  
4325 4330 4335

Val Asp Lys Val Arg Ala Val Ile Ala Asp Arg Thr Glu Val Thr Gly  
4340 4345 4350

Val Leu Ser Leu Leu Ala Met Asp Gly Met Pro Cys Ala Ala His Pro  
4355 4360 4365

His Leu Ser Arg Gl Val Ala Ala Thr Val Ile Leu Th Gln Val Leu  
 4370 4375 4380

Gly Asp Ala Gly Val Ser Ala Pro Leu Trp Leu Ala Thr Thr Gly Gly  
 385 4390 4395 4400

Val Glu Ala Gly Thr Glu Asp Gly Pro Ala Asp Pro Asp His Gly Leu  
 4405 4410 4415

Ile Trp Gly Leu Gly Arg Val Val Gly Leu Glu His Pro Gln Trp Trp  
 4420 4425 4430

Gly Gly Leu Ile Asp Leu Pro Glu Thr Leu Asp Glu Thr Ser Arg Asn  
 4435 4440 4445

Gly Leu Val Ala Ala Leu Ala Gly Thr Ala Ala Glu Asp Gln Leu Ala  
 4450 4455 4460

Val Arg Ser Ser Gly Leu Phe Val Arg Arg Val Val Arg Ala Ala Arg  
 4465 4470 4475 4480

Asn Pro Arg Ser Glu Thr Trp Arg Ser Arg Gly Thr Val Leu Ile Thr  
 4485 4490 4495

Gly Gly Thr Gly Ala Leu Gly Ala Glu Val Ala Arg Trp Leu Ala Arg  
 4500 4505 4510

Arg Gly Ala Glu His Leu Val Leu Ile Ser Arg Arg Gly Pro Glu Ala  
 4515 4520 4525

Pro Gly Ala Ala Asp Leu Gly Ala Glu Leu Thr Glu Leu Gly Val Lys  
 4530 4535 4540

Val Thr Val Leu Ala Cys Asp Val Thr Asp Arg Asp Glu Leu Ala Ala  
 4545 4550 4555 4560

Val Leu Ala Ala Val Pro Thr Glu Tyr Pro Leu Ser Ala Val Val His  
 4565 4570 4575

Thr Ala Gly Val Gly Thr Pro Ala Asn Leu Ala Glu Thr Thr Leu Ala  
 4580 4585 4590

Gln Phe Ala Asp Val Leu Ser Ala Lys Val Val Gly Ala Ala Asn Leu  
 4595 4600 4605

Asp Arg Leu Leu Gly Gly Gln Pro Leu Asp Ala Phe Val Leu Phe Ser  
 4610 4615 4620



Ser Ile Ser Gly Val Trp Gly Ala Gly Gly Gln Gly Ala Tyr Ser Ala  
 625 4630 4635 4640

Ala Asn Ala Tyr Leu Asp Ala Leu Ala Glu Arg Arg Arg Ala Cys Gly  
 4645 4650 4655

Arg Pro Ala Thr Cys Ile Ala Trp Gly Pro Trp Ala Gly Ala Gly Met  
 4660 4665 4670

Ala Val Gln Glu Gly Asn Glu Ala His Leu Arg Arg Arg Gly Leu Val  
 4675 4680 4685

Pro Met Glu Pro Gln Ser Ala Leu Phe Ala Leu Gln Gln Ala Leu Ser  
 4690 4695 4700

Gln Arg Glu Thr Ala Ile Thr Val Ala Asp Val Asp Trp Glu Arg Phe  
 705 4710 4715 4720

Ala Ala Ser Phe Thr Ala Ala Arg Pro Arg Pro Leu Leu Glu Glu Ile  
 4725 4730 4735

Val Asp Leu Arg Pro Asp Thr Glu Thr Glu Glu Lys His Gly Ala Gly  
 4740 4745 4750

Glu Leu Gly Gln Gln Leu Ala Ala Leu Pro Pro Ala Glu Arg Gly His  
 4755 4760 4765

Leu Leu Leu Glu Val Val Leu Ala Glu Thr Ala Ser Thr Leu Gly His  
 4770 4775 4780

Asp Ser Ala Glu Ala Val Gln Pro Asp Arg Thr Phe Ala Glu Leu Gly  
 785 4790 4795 4800

Phe Asp Ser Leu Thr Ala Val Glu Leu Arg Asn Arg Leu Asn Ala Val  
 4805 4810 4815

Thr Gly Leu Arg Leu Pro Pro Thr Leu Val Phe Asp His Pro Thr Pro  
 4820 4825 4830

Leu Ala Leu Ser Glu Gln Leu Val Pro Ala Leu Val Ala Glu Pro Asp  
 4835 4840 4845

Asn Gly Ile Glu Ser Leu Leu Ala Glu Leu Asp Arg Leu Asp Thr Thr  
 4850 4855 4860

Leu Ala Gln Gly Pro Ser Ile Pro Leu Glu Asp Gln Ala Lys Val Ala  
 865 4870 4875 4880

Glu Arg Leu His A Leu Leu Ala Lys Trp Asp Gly A Arg Asp Gly  
 4885 4890 4895

Thr Ala Arg Ala Thr Ser Pro Gln Ser Leu Thr Ala Ala Thr Asp Asp  
 4900 4905 4910

Glu Ile Phe Asp Leu Ile Asp Arg Lys Phe Arg Arg  
 4915 4920

<210> 49  
 <211> 16767  
 <212> DNA  
 <213> Saccharopolyspora spinosa

<220>  
 <221> CDS  
 <222> (1)..(16764)  
 <223> ORF22; Polyketidsynthase

<400> 49  
 atg gcc aat gaa gaa aag ctc cgc gag tac ctc aag cgt gtc gtc gtc 48  
 Met Ala Asn Glu Glu Lys Leu Arg Glu Tyr Leu Lys Arg Val Val Val  
 1 5 10 15  
 gaa ctg gaa gag gcg cac gaa cgc ctg cac gag ttg gag cgc cag gag 96  
 Glu Leu Glu Glu Ala His Glu Arg Leu His Glu Leu Glu Arg Gln Glu  
 20 25 30  
 cac gac ccc atc gcg atc gtg tcg atg gga tgt cgt tat ccc ggt ggc 144  
 His Asp Pro Ile Ala Ile Val Ser Met Gly Cys Arg Tyr Pro Gly Gly  
 35 40 45  
 gtc tcc act ccg gag gag ctg tgg cga ctg gtc gtc gac gga gga gac 192  
 Val Ser Thr Pro Glu Glu Leu Trp Arg Leu Val Val Asp Gly Gly Asp  
 50 55 60  
 gcg atc gcg aac ttc ccc gaa gac cgt ggc tgg aat ctg gac gag ctg 240  
 Ala Ile Ala Asn Phe Pro Glu Asp Arg Gly Trp Asn Leu Asp Glu Leu  
 65 70 75 80  
 ttc gat cct gat ccg ggc cga gcc ggg acc tcc tac gtc cgc gag ggt 288  
 Phe Asp Pro Asp Pro Gly Arg Ala Gly Thr Ser Tyr Val Arg Glu Gly  
 85 90 95  
 ggt ttc ctg cgc ggg gtc gcg gac ttc gat gcc ggg ctc ttc ggg atc 336  
 Gly Phe Leu Arg Gly Val Ala Asp Phe Asp Ala Gly Leu Phe Gly Ile

100	105	110	
agt ccg cgc gag gca cag gcg atg gac ccg caa cag cgg ttg ctg ctg			384
Ser Pro Arg Glu Ala Gln Ala Met Asp Pro Gln Gln Arg Leu Leu Leu			
115	120	125	
gag atc tcg tgg gag gtg ttc gag cgc gcc ggc att gac ccg ttt tct			432
Glu Ile Ser Trp Glu Val Phe Glu Arg Ala Gly Ile Asp Pro Phe Ser			
130	135	140	
ttg cgg ggt acc aag acc ggt gtg ttc gcg ggc ctg atc tac cac gac			480
Leu Arg Gly Thr Lys Thr Gly Val Phe Ala Gly Leu Ile Tyr His Asp			
145	150	155	160
tac gcg tcg cgg ttt cgc aag acc ccc gcg gag ttc gag ggt tac ttc			528
Tyr Ala Ser Arg Phe Arg Lys Thr Pro Ala Glu Phe Glu Gly Tyr Phe			
165	170	175	
gcc acc ggc aac gcg ggc agc gtc gca tcc ggc cgg gtg gct tac acc			576
Ala Thr Gly Asn Ala Gly Ser Val Ala Ser Gly Arg Val Ala Tyr Thr			
180	185	190	
ttc ggg tta gag ggc ccg gcg gtc acc gtg gac acc gcc tgc tcg tcg			624
Phe Gly Leu Glu Gly Pro Ala Val Thr Val Asp Thr Ala Cys Ser Ser			
195	200	205	
tcc ctg gtg gcg ctg cac ctg gcc tgc cag tcc ctg cgg ctg ggc gaa			672
Ser Leu Val Ala Leu His Leu Ala Cys Gln Ser Leu Arg Leu Gly Glu			
210	215	220	
tgc gac ctg gcc ctg gcc ggt ggc att tcg gtg atg gcc acg ccg gga			720
Cys Asp Leu Ala Leu Ala Gly Gly Ile Ser Val Met Ala Thr Pro Gly			
225	230	235	240
gcc ttc gtc gag ttc agc cgg caa cgc gca ctc gcc tcg gat ggc cgg			768
Ala Phe Val Glu Phe Ser Arg Gln Arg Ala Leu Ala Ser Asp Gly Arg			
245	250	255	
tgc aag ccc ttc gcg gat gcc gcc gac ggc acc ggc tgg ggc gag ggc			816
Cys Lys Pro Phe Ala Asp Ala Ala Asp Gly Thr Gly Trp Gly Glu Gly			
260	265	270	
gcc gga atg ctg ctg ctg gaa cgg ctg tcg gac gca cga cga aac ggc			864
Ala Gly Met Leu Leu Leu Glu Arg Leu Ser Asp Ala Arg Arg Asn Gly			
275	280	285	
cac ccg gtg ctg gcg gcg gtg gtc ggt tcc gcg atc aac cag gac ggg			912
His Pro Val Leu Ala Ala Val Val Gly Ser Ala Ile Asn Gln Asp Gly			

290	295	300	
acg tcc aac ggc ctg acc gcg ccc agc ggt ccc gca cag cag cga gtg			960
Thr Ser Asn Gly Leu Thr Ala Pro Ser Gly Pro Ala Gln Gln Arg Val			
305	310	315	320
atc cgc caa gcc ctg gcg aac gcc ggg ttg tcg ccc gcc gag gtc gat			1008
Ile Arg Gln Ala Leu Ala Asn Ala Gly Leu Ser Pro Ala Glu Val Asp			
325	330		335
gtg gtc gag gcg cac ggc acg ggc acg gcc ttg ggc gac ccg atc gag			1056
Val Val Glu Ala His Gly Thr Gly Thr Ala Leu Gly Asp Pro Ile Glu			
340	345		350
gcg cag gcc ctg atc gcc acc tac ggg gcg aac cgg tcg gcg gat cat			1104
Ala Gln Ala Leu Ile Ala Thr Tyr Gly Ala Asn Arg Ser Ala Asp His			
355	360		365
ccg ctg ctg ctg ggt tcc ctc aag tcg aac atc ggc cac acc cag gct			1152
Pro Leu Leu Leu Gly Ser Leu Lys Ser Asn Ile Gly His Thr Gln Ala			
370	375		380
gcc gcc ggt gtg gcc ggg gtg atc aag tcg gtc ctg gcc atc agg cac			1200
Ala Ala Gly Val Ala Gly Val Ile Lys Ser Val Leu Ala Ile Arg His			
385	390	395	400
cgg gag atg ccc cgc agc ctg cac atc gac cag cca tcg cag cac gtg			1248
Arg Glu Met Pro Arg Ser Leu His Ile Asp Gln Pro Ser Gln His Val			
405	410		415
gac tgg tcg gcg ggc gcg gtg cgg ctg ctc acg gac agc gtt gac tgg			1296
Asp Trp Ser Ala Gly Ala Val Arg Leu Leu Thr Asp Ser Val Asp Trp			
420	425		430
ccg gat ctc ggc agg ccg cgc cga gca ggg gtg tcc tcg ttc ggc atg			1344
Pro Asp Leu Gly Arg Pro Arg Arg Ala Gly Val Ser Ser Phe Gly Met			
435	440		445
agc ggt acc aac gca cac ctg atc gtc gag gaa gta tcc gac gag ccg			1392
Ser Gly Thr Asn Ala His Leu Ile Val Glu Glu Val Ser Asp Glu Pro			
450	455		460
gtc tcg ggc agt acc gag ccg acc ggg gca ttt ccc tgg ccg ctg tcc			1440
Val Ser Gly Ser Thr Glu Pro Thr Gly Ala Phe Pro Trp Pro Leu Ser			
465	470	475	480
ggc aag acg gag acg gca ttg cgc gag cag gct gcc gag ttg ctc tcc			1488
Gly Lys Thr Glu Thr Ala Leu Arg Glu Gln Ala Ala Glu Leu Leu Ser			

485	490	495	
gta gtg acc gag cac ccg gag ccg gga ctg ggg gac gtc ggg tac tcg			1536
Val Val Thr Glu His Pro Glu Pro Gly Leu Gly Asp Val Gly Tyr Ser			
500	505	510	
ctg gcc acc ggt cgc gct gcg atg gag cac cgg gct gtc gtg gtt gcc			1584
Leu Ala Thr Gly Arg Ala Ala Met Glu His Arg Ala Val Val Val Ala			
515	520	525	
gac gat cgg gac tct ttc gtc gcc gga ctg acg gcg ttg gct gcg ggc			1632
Asp Asp Arg Asp Ser Phe Val Ala Gly Leu Thr Ala Leu Ala Ala Gly			
530	535	540	
gtt ccg gca gcc aac gtg gtg cag ggc gcg gcc gac tgc aag gga aag			1680
Val Pro Ala Ala Asn Val Val Gln Gly Ala Ala Asp Cys Lys Gly Lys			
545	550	555	560
gtc gcg ttc gtg ttc ccc ggc cag ggc tcg cat tgg cag ggg atg gcg			1728
Val Ala Phe Val Phe Pro Gly Gln Gly Ser His Trp Gln Gly Met Ala			
565	570	575	
agg gaa ctg tcc gaa tcc tcg ccg gtg ttc cgg cgg aag ctg gcg gaa			1776
Arg Glu Leu Ser Glu Ser Ser Pro Val Phe Arg Arg Lys Leu Ala Glu			
580	585	590	
tgc gcg gcg gct acg gcc cct tac gtg gac tgg tcg ctg ctc ggc gtc			1824
Cys Ala Ala Ala Thr Ala Pro Tyr Val Asp Trp Ser Leu Leu Gly Val			
595	600	605	
ctt cgc ggt gat ccc gat gca ccc gcg ctg gat cgc gac gac gtg att			1872
Leu Arg Gly Asp Pro Asp Ala Pro Ala Leu Asp Arg Asp Asp Val Ile			
610	615	620	
cag ctc gcg ctg ttc gcc atg atg gtg tcg ctg gcc gaa ctg tgg cgt			1920
Gln Leu Ala Leu Phe Ala Met Met Val Ser Leu Ala Glu Leu Trp Arg			
625	630	635	640
tcg tgc gga gtg gag ccc gcc gcg gtg gtc ggt cat tcc cag ggc gag			1968
Ser Cys Gly Val Glu Pro Ala Ala Val Val Gly His Ser Gln Gly Glu			
645	650	655	
atc gcc gcc gcc cat gtg gca ggc gct ttg tcc ttg act gat gcg gtg			2016
Ile Ala Ala Ala His Val Ala Gly Ala Leu Ser Leu Thr Asp Ala Val			
660	665	670	
cgc atc atc gct gcc cgc tgc gat gcg gtg tcg gcg ctg acc ggg aag			2064
Arg Ile Ile Ala Ala Arg Cys Asp Ala Val Ser Ala Leu Thr Gly Lys			

675

680

685

gga ggc atg ctc gcg att gcc ttg ccg gaa agc gcg gtg gtg aag cga 2112  
 Gly Gly Met Leu Ala Ile Ala Leu Pro Glu Ser Ala Val Val Lys Arg  
 690 695 700

atc gca ggc ctg ccg gag ctg acc gtt gcg gcg gtc aac gga ccc ggc 2160  
 Ile Ala Gly Leu Pro Glu Leu Thr Val Ala Ala Val Asn Gly Pro Gly  
 705 710 715 720

tcc act gtc gtt tcc ggc gaa ccg tcg gct ctg gag cgt ctg cag acc 2208  
 Ser Thr Val Val Ser Gly Glu Pro Ser Ala Leu Glu Arg Leu Gln Thr  
 725 730 735

gaa ctg acc gcg gaa aac gtg cag acc ccg ccg gtg gga att gat tac 2256  
 Glu Leu Thr Ala Glu Asn Val Gln Thr Arg Arg Val Gly Ile Asp Tyr  
 740 745 750

gcc tcg cat tcg ccg cag atc gcg cag gtc cag ggc ccg ctt ctg gac 2304  
 Ala Ser His Ser Pro Gln Ile Ala Gln Val Gln Gly Arg Leu Leu Asp  
 755 760 765

cgg ctg ggc gaa gtc ggg tcc gaa cct gct gag atc gct ttc tac tcg 2352  
 Arg Leu Gly Glu Val Gly Ser Glu Pro Ala Glu Ile Ala Phe Tyr Ser  
 770 775 780

acg gtc acc ggc gag ccg acg gac acc ggc cga ctc gac gcc gac tac 2400  
 Thr Val Thr Gly Glu Arg Thr Asp Thr Gly Arg Leu Asp Ala Asp Tyr  
 785 790 795 800

tgg tac cag aac ctt ccg cag ccc gtc cgc ttc cag cag acc gtc gcc 2448  
 Trp Tyr Gln Asn Leu Arg Gln Pro Val Arg Phe Gln Gln Thr Val Ala  
 805 810 815

cgg atg gca gat cag ggc tat ccg ttc ttc gtc gag gtg agc ccg cac 2496  
 Arg Met Ala Asp Gln Gly Tyr Arg Phe Phe Val Glu Val Ser Pro His  
 820 825 830

ccg ctg ctc acc gcc gga atc cag gaa acg ctg gaa gcc gcg gac gcg 2544  
 Pro Leu Leu Thr Ala Gly Ile Gln Glu Thr Leu Glu Ala Ala Asp Ala  
 835 840 845

ggc ggg gtg gtg gtc ggt tcg ctg ccg cgt gcc gag gcc gcc tcc ccg 2592  
 Gly Gly Val Val Val Gly Ser Leu Arg Arg Gly Glu Gly Gly Ser Arg  
 850 855 860

cgc tgg ctg act tcg ctg gcc gag tgc cag gtg cgc gga ctg ccg gtg 2640  
 Arg Trp Leu Thr Ser Leu Ala Glu Cys Gln Val Arg Gly Leu Pro Val

865	870	875	880	
aat tgg gaa cag gta ttc ctc aac acc gga gcc cga cgc gtg ccg ctg				2688
Asn Trp Glu Gln Val Phe Leu Asn Thr Gly Ala Arg Arg Val Pro Leu	885	890	895	
ccg acc tac ccg ttc cag cgg cag cgg tac tgg ttg gag tcc gcc gag				2736
Pro Thr Tyr Pro Phe Gln Arg Gln Arg Tyr Trp Leu Glu Ser Ala Glu	900	905	910	
tac gac gcg ggc gat ctc ggt tgc gtg ggc ttg ctc tcc gcc gag cat				2784
Tyr Asp Ala Gly Asp Leu Gly Ser Val Gly Leu Leu Ser Ala Glu His	915	920	925	
ccc ctg ctc ggg gct gcg gtg acg ctg gcc gat gcg ggc ggg ttc ctg				2832
Pro Leu Leu Gly Ala Ala Val Thr Leu Ala Asp Ala Gly Gly Phe Leu	930	935	940	
ctg acc ggc aag ctg tgc gtc aag acc cag ccc tgg ttg gcc gac cac				2880
Leu Thr Gly Lys Leu Ser Val Lys Thr Gln Pro Trp Leu Ala Asp His	945	950	955	960
gtg gtc ggc ggg gcg atc ctg ctg ccc ggc acc gcg ttc gtg gaa atg				2928
Val Val Gly Gly Ala Ile Leu Leu Pro Gly Thr Ala Phe Val Glu Met	965	970	975	
ctg ata cgc gcc gcg gac cag gtc ggg tgc gat ctg atc gag gag ttg				2976
Leu Ile Arg Ala Ala Asp Gln Val Gly Cys Asp Leu Ile Glu Glu Leu	980	985	990	
tcc ctg acg act ccg ctg gtt ttg ccc gcg acc ggt gcg gtg cag gtg				3024
Ser Leu Thr Thr Pro Leu Val Leu Pro Ala Thr Gly Ala Val Gln Val	995	1000	1005	
cag atc gcg gtt ggc ggt ccg gac gag gcc ggg cgc cgc tgc gtc cgc				3072
Gln Ile Ala Val Gly Gly Pro Asp Glu Ala Gly Arg Arg Ser Val Arg	1010	1015	1020	
gtg cat tcc tgt cga gac gac gcc gtg ccg cag gac tgc tgg acc tgc				3120
Val His Ser Cys Arg Asp Asp Ala Val Pro Gln Asp Ser Trp Thr Cys	1025	1030	1035	1040
cac gcg acc ggc acg ttg acc tcc agc gat cac cag gac gcc ggc cag				3168
His Ala Thr Gly Thr Leu Thr Ser Ser Asp His Gln Asp Ala Gly Gln	1045	1050	1055	
ggc ccc gat ggg att tgg ccg ccc aac gat gct gtc gcg gtt ccg ctg				3216
Gly Pro Asp Gly Ile Trp Pro Pro Asn Asp Ala Val Ala Val Pro Leu				

1060

1065

1070

gac agc ttc tac gcc cgc gca gct gag cgg ggc ttc gat ttc ggc ccg 3264  
 Asp Ser Phe Tyr Ala Arg Ala Ala Glu Arg Gly Phe Asp Phe Gly Pro  
 1075 1080 1085

gcg ttc cag ggg ttg cag gcg gct tgg aag cgc gga gac gag atc ttc 3312  
 Ala Phe Gln Gly Leu Gln Ala Ala Trp Lys Arg Gly Asp Glu Ile Phe  
 1090 1095 1100

gcc gag gtc ggc ctg ccc acc gca cac cgc gaa gac gcc ggc agg ttc 3360  
 Ala Glu Val Gly Leu Pro Thr Ala His Arg Glu Asp Ala Gly Arg Phe  
 1105 1110 1115 1120

gga atc cac cct gct ctg ctg gat gcg gca ctg cag gcg ctg ggc gca 3408  
 Gly Ile His Pro Ala Leu Leu Asp Ala Ala Leu Gln Ala Leu Gly Ala  
 1125 1130 1135

gcc gaa gag gat ccg gac gag gga tgg ctc ccg ttc gcg tgg caa ggt 3456  
 Ala Glu Glu Asp Pro Asp Glu Gly Trp Leu Pro Phe Ala Trp Gln Gly  
 1140 1145 1150

gtg tcc ctc aaa gcg acg ggc gca ctt tcc ctt cgg gtg cac ctc gtt 3504  
 Val Ser Leu Lys Ala Thr Gly Ala Leu Ser Leu Arg Val His Leu Val  
 1155 1160 1165

ccg gcg ggc gcg aat gcg gtg tcg gtg ttc acg acc gac acg act ggc 3552  
 Pro Ala Gly Ala Asn Ala Val Ser Val Phe Thr Thr Asp Thr Thr Gly  
 1170 1175 1180

caa gcc gtg ctc tcc atc gat tcg ctg gtg ctg cgc cag att tcg gac 3600  
 Gln Ala Val Leu Ser Ile Asp Ser Leu Val Leu Arg Gln Ile Ser Asp  
 1185 1190 1195 1200

aag cag ttg gca gcg gcc cgt gcg atg gaa cac gag tcc ctg ttc ccg 3648  
 Lys Gln Leu Ala Ala Ala Arg Ala Met Glu His Glu Ser Leu Phe Arg  
 1205 1210 1215

gtc gac tgg aag cga atc tcg ccc ggc gct gcc aag ccg gtc tcc tgg 3696  
 Val Asp Trp Lys Arg Ile Ser Pro Gly Ala Ala Lys Pro Val Ser Trp  
 1220 1225 1230

gca gtg atc ggc aat gac gaa ctc gcc cga gcc tgc ggc tcg gca ctt 3744  
 Ala Val Ile Gly Asn Asp Glu Leu Ala Arg Ala Cys Gly Ser Ala Leu  
 1235 1240 1245

ggc acg gaa ctc cac ccc gac ctg acc ggg ttg gct gac ccg ccc ccg 3792  
 Gly Thr Glu Leu His Pro Asp Leu Thr Gly Leu Ala Asp Pro Pro Pro



1250	1255	1260	
gac gtc gtg gtg gtg cca tgc ggt gcg tct cgc cag gac ttg gac gtt			3840
Asp Val Val Val Val Pro Cys Gly Ala Ser Arg Gln Asp Leu Asp Val			
1265	1270	1275	1280
gct tcc gag gca cgt gcc gcg aca caa cgc atg ctt gac ctg atc cag			3888
Ala Ser Glu Ala Arg Ala Ala Thr Gln Arg Met Leu Asp Leu Ile Gln			
	1285	1290	1295
gat tgg ttg gcg gcg gcg cga ttc gcc gga tct cgc ctg gtg gtt gtg			3936
Asp Trp Leu Ala Ala Ala Arg Phe Ala Gly Ser Arg Leu Val Val Val			
	1300	1305	1310
acg tgt ggt gcg gcg tcg aca ggt ccc gcc gag ggt gtt tcc gac ctg			3984
Thr Cys Gly Ala Ala Ser Thr Gly Pro Ala Glu Gly Val Ser Asp Leu			
	1315	1320	1325
gtg cat gct gcg tcg tgg ggt ttg ttg cgt tcg gcg cag tcg gag aac			4032
Val His Ala Ala Ser Trp Gly Leu Leu Arg Ser Ala Gln Ser Glu Asn			
	1330	1335	1340
ccg gac cga ttc gtg ttg gtc gat gtg gac gga acc gcc gaa tca tgg			4080
Pro Asp Arg Phe Val Leu Val Asp Val Asp Gly Thr Ala Glu Ser Trp			
	1345	1350	1355
			1360
cgt gcg ctc gcg gcg gcc gtg cgt tcc gga gaa ccg cag ctg gcg ttg			4128
Arg Ala Leu Ala Ala Ala Val Arg Ser Gly Glu Pro Gln Leu Ala Leu			
	1365	1370	1375
cgc gcc ggt gaa gtc cgg gtg cct cgc ctg gcg cga tgt gtt gcc gcc			4176
Arg Ala Gly Glu Val Arg Val Pro Arg Leu Ala Arg Cys Val Ala Ala			
	1380	1385	1390
gag gac agc cgg atc cca gtg ccc ggt gcg gat ggg acg gtg ttg att			4224
Glu Asp Ser Arg Ile Pro Val Pro Gly Ala Asp Gly Thr Val Leu Ile			
	1395	1400	1405
tcc ggc ggt acg ggc ctg ctg ggc ggg ttg gtt gcc cgg cat ttg gtg			4272
Ser Gly Gly Thr Gly Leu Leu Gly Gly Leu Val Ala Arg His Leu Val			
	1410	1415	1420
gcg gag cgc ggt gtc cgc cgc ctg gtg ctc gcg ggg cga cgc ggc tgg			4320
Ala Glu Arg Gly Val Arg Arg Leu Val Leu Ala Gly Arg Arg Gly Trp			
	1425	1430	1435
			1440
agc gcc ccc ggg gtc acc gac ctg gtg gat gag ttg gtg ggc ctg gga			4368
Ser Ala Pro Gly Val Thr Asp Leu Val Asp Glu Leu Val Gly Leu Gly			

1450

1450

1455

gct gcg gtc gag gtg gcg agc tgc gat gtc ggg gat cgg gcc cag ttg 4416  
 Ala Ala Val Glu Val Ala Ser Cys Asp Val Gly Asp Arg Ala Gln Leu  
 1460 1465 1470

gac cgg ctg ctg acg acg atc tgc gca gag ttc ccg ctg cgc gga gtg 4464  
 Asp Arg Leu Leu Thr Thr Ile Ser Ala Glu Phe Pro Leu Arg Gly Val  
 1475 1480 1485

gtg cat gcg gcc ggg gca ctt gcc gac ggg gtc gtc gag tgc ctg aca 4512  
 Val His Ala Ala Gly Ala Leu Ala Asp Gly Val Val Glu Ser Leu Thr  
 1490 1495 1500

cca gag cac gtg gca aag gtg ttc ggc ccg aag gcc gcc ggt gcg tgg 4560  
 Pro Glu His Val Ala Lys Val Phe Gly Pro Lys Ala Ala Gly Ala Trp  
 1505 1510 1515 1520

cac ctg cac gag ttg act ctt gat ctg gat ctc tgc ttc ttc gtg ctc 4608  
 His Leu His Glu Leu Thr Leu Asp Leu Asp Leu Ser Phe Phe Val Leu  
 1525 1530 1535

ttc tcc tgc ttc tcc ggc gtg gcg ggg gct gcg ggt cag gga aac tac 4656  
 Phe Ser Ser Phe Ser Gly Val Ala Gly Ala Ala Gly Gln Gly Asn Tyr  
 1540 1545 1550

gcg gcg gcg aac gcg ttc ctg gac ggc ctg gct cag cac cgg cgg acg 4704  
 Ala Ala Ala Asn Ala Phe Leu Asp Gly Leu Ala Gln His Arg Arg Thr  
 1555 1560 1565

gcg ggg ctg cct gcg gtg tgc ctg gct tgg ggc ttg tgg gag cag ccc 4752  
 Ala Gly Leu Pro Ala Val Ser Leu Ala Trp Gly Leu Trp Glu Gln Pro  
 1570 1575 1580

agc ggg atg acc gga gcg ctc gat gcg gcg ggc cgt agc cgc att gcg 4800  
 Ser Gly Met Thr Gly Ala Leu Asp Ala Ala Gly Arg Ser Arg Ile Ala  
 1585 1590 1595 1600

cgc acc aat ccg ccg atg tcc gcg ccg gac ggg ttg cgg ctg ttc gag 4848  
 Arg Thr Asn Pro Pro Met Ser Ala Pro Asp Gly Leu Arg Leu Phe Glu  
 1605 1610 1615

atg gcg ttt cgc gtt ccg ggc gaa tgc ctt ctg gtt ccg gtc cac gtc 4896  
 Met Ala Phe Arg Val Pro Gly Glu Ser Leu Leu Val Pro Val His Val  
 1620 1625 1630

gac ctg aac gcc ctg cgc gct gat gcg gcc gac ggc ggt gtg cct gcg 4944  
 Asp Leu Asn Ala Leu Arg Ala Asp Ala Ala Asp Gly Gly Val Pro Ala

1635	1640	1645	
ttg ttg cgc gac ctg gtg cca gcg ccc gtg cgg cgg agc gcg gtc aac			4992
Leu Leu Arg Asp Leu Val Pro Ala Pro Val Arg Arg Ser Ala Val Asn			
1650	1655	1660	
gag tcg gcg gac gtc aac ggt ctg gtt ggt cgg ctg cgg agg ctg ccg			5040
Glu Ser Ala Asp Val Asn Gly Leu Val Gly Arg Leu Arg Arg Leu Pro			
1665	1670	1675	1680
gac ctg gat cag gaa acc cag ctg ttg ggt ttg gtg cgc gag cat gtt			5088
Asp Leu Asp Gln Glu Thr Gln Leu Leu Gly Leu Val Arg Glu His Val			
1685	1690	1695	
tcg gcg gtg ctg ggg cat tcg ggt gcg gtc gag gtc ggg gcc gat cgt			5136
Ser Ala Val Leu Gly His Ser Gly Ala Val Glu Val Gly Ala Asp Arg			
1700	1705	1710	
gct ttc cgg gat ttg ggt ttt gat tcg ttg tcc ggt gtg gag ttt cgg			5184
Ala Phe Arg Asp Leu Gly Phe Asp Ser Leu Ser Gly Val Glu Phe Arg			
1715	1720	1725	
aac cgg ctt ggc ggg gtg ctg ggc gtt cgg ttg ccg gct act gcg gtg			5232
Asn Arg Leu Gly Gly Val Leu Gly Val Arg Leu Pro Ala Thr Ala Val			
1730	1735	1740	
ttc gac tat ccg aca ccg cgg gcg ttg gtt cgg ttc ttg ctc gac aaa			5280
Phe Asp Tyr Pro Thr Pro Arg Ala Leu Val Arg Phe Leu Leu Asp Lys			
1745	1750	1755	1760
ctg att ggt ggc gtg gag gct ccg act ccc gca ccg gcg gct gtg gcg			5328
Leu Ile Gly Gly Val Glu Ala Pro Thr Pro Ala Pro Ala Ala Val Ala			
1765	1770	1775	
gcg gtg act gct gac gat ccc gtt gtg atc gtg ggg atg ggc tgt cgt			5376
Ala Val Thr Ala Asp Asp Pro Val Val Ile Val Gly Met Gly Cys Arg			
1780	1785	1790	
tat ccg ggt ggg gtg tcc tcg ccg gag gag ctt tgg cgt ttg gtg gcc			5424
Tyr Pro Gly Gly Val Ser Ser Pro Glu Glu Leu Trp Arg Leu Val Ala			
1795	1800	1805	
ggg ggc ttg gat gcg gtg gcg gag ttc ccg gac gat cgt ggc tgg gat			5472
Gly Gly Leu Asp Ala Val Ala Glu Phe Pro Asp Asp Arg Gly Trp Asp			
1810	1815	1820	
cag gcg ggg ttg ttc gat ccg gat ccc gat cgt ctt ggg acc tcg tat			5520
Gln Ala Gly Leu Phe Asp Pro Asp Pro Asp Arg Leu Gly Thr Ser Tyr			

1825	30	1835	1840
gtg tgt gag ggt ggc ttc ctg cga gat gcg gca gag ttc gat gcc ggt			5568
Val Cys Glu Gly Gly Phe Leu Arg Asp Ala Ala Glu Phe Asp Ala Gly			
1845	1850	1855	
ttc ttc ggg att tcc ccg cgt gag gcg ttg gcg atg gat ccg cag cag			5616
Phe Phe Gly Ile Ser Pro Arg Glu Ala Leu Ala Met Asp Pro Gln Gln			
1860	1865	1870	
cgg ttg ctg ctg gaa gtc gct tgg gaa acc gtg gag cgg gcg ggg att			5664
Arg Leu Leu Leu Glu Val Ala Trp Glu Thr Val Glu Arg Ala Gly Ile			
1875	1880	1885	
gat ccg ctt tcg ttg cgg ggg agc cgg acc ggc gtg ttc gcg ggg ctg			5712
Asp Pro Leu Ser Leu Arg Gly Ser Arg Thr Gly Val Phe Ala Gly Leu			
1890	1895	1900	
atg cac cac gac tac ggc gcg cgg ttc atc acg agg gcg ccg gag ggt			5760
Met His His Asp Tyr Gly Ala Arg Phe Ile Thr Arg Ala Pro Glu Gly			
1905	1910	1915	1920
ttc gag ggt tat cta ggt aat ggc agc gcg gga ggc gtg ttt tcg ggt			5808
Phe Glu Gly Tyr Leu Gly Asn Gly Ser Ala Gly Gly Val Phe Ser Gly			
1925	1930	1935	
cgg gtt gcg tat tcg ttt ggt ttc gag ggt cct gcg gtg acg gtg gat			5856
Arg Val Ala Tyr Ser Phe Gly Phe Glu Gly Pro Ala Val Thr Val Asp			
1940	1945	1950	
acg gcg tgt tcg tcg tcg ttg gtg gcg ctg cac ctg gcg ggt caa gca			5904
Thr Ala Cys Ser Ser Ser Leu Val Ala Leu His Leu Ala Gly Gln Ala			
1955	1960	1965	
ctg cgg tct ggt gag tgt gat ctg gct ctt gcg ggt ggt gtg acg gtg			5952
Leu Arg Ser Gly Glu Cys Asp Leu Ala Leu Ala Gly Gly Val Thr Val			
1970	1975	1980	
atg gcc acg ccg ggg atg ttc gtg gag ttt tcg cgt caa cgg ggc ttg			6000
Met Ala Thr Pro Gly Met Phe Val Glu Phe Ser Arg Gln Arg Gly Leu			
1985	1990	1995	2000
gcg gcg gat ggg cgg tgc aag tcg ttt gcg gcg gct gcg gat ggc acc			6048
Ala Ala Asp Gly Arg Cys Lys Ser Phe Ala Ala Ala Ala Asp Gly Thr			
2005	2010	2015	
ggt tgg gga gaa ggc gcg ggc ttg gtg ttg ttg gag cgg ctg tcg gat			6096
Gly Trp Gly Glu Gly Ala Gly Leu Val Leu Leu Glu Arg Leu Ser Asp			

2020	2025	2030	
gcc cgg cgc aac ggg cac gcg gtt ctg gcg gtc gtg cgg ggt agc gcg Ala Arg Arg Asn Gly His Ala Val Leu Ala Val Val Arg Gly Ser Ala 2035 2040 2045			6144
gtg aat cag gat ggt gcg tcg aat ggt ttg acg gcg ccg aat ggg ccc Val Asn Gln Asp Gly Ala Ser Asn Gly Leu Thr Ala Pro Asn Gly Pro 2050 2055 2060			6192
tcg cag cag cgg gtg atc acg cag gcg ttg gcg agt gct ggt ttg tcg Ser Gln Gln Arg Val Ile Thr Gln Ala Leu Ala Ser Ala Gly Leu Ser 2065 2070 2075 2080			6240
gtg tct gat gtg gac gcc gtg gag gcg cat ggg act gga acc agg ctt Val Ser Asp Val Asp Ala Val Glu Ala His Gly Thr Gly Thr Arg Leu 2085 2090 2095			6288
ggg gat ccg att gag gcg cag gct ctg att gcc act tac ggg cag ggg Gly Asp Pro Ile Glu Ala Gln Ala Leu Ile Ala Thr Tyr Gly Gln Gly 2100 2105 2110			6336
cgg gat agc gat cgg ccg ttg tgg ttg ggg tcg gtg aag tcg aat att Arg Asp Ser Asp Arg Pro Leu Trp Leu Gly Ser Val Lys Ser Asn Ile 2115 2120 2125			6384
ggg cat acg cag gcg gcg gcg ggt gtc gct ggt gtg atc aag atg gtg Gly His Thr Gln Ala Ala Ala Gly Val Ala Gly Val Ile Lys Met Val 2130 2135 2140			6432
atg gcg atg cgg cac ggg cag ctg ccc gcg acg ttg cat gtg gat gaa Met Ala Met Arg His Gly Gln Leu Pro Ala Thr Leu His Val Asp Glu 2145 2150 2155 2160			6480
cct acg tcg gaa gtg gat tgg tcg gcg ggg gat gtc cag ctc ctc acg Pro Thr Ser Glu Val Asp Trp Ser Ala Gly Asp Val Gln Leu Leu Thr 2165 2170 2175			6528
gag aac acc ccc tgg ccc ggc aac agc cat cct cgg cgg gtg ggc gtg Glu Asn Thr Pro Trp Pro Gly Asn Ser His Pro Arg Arg Val Gly Val 2180 2185 2190			6576
tcg tcg ttc ggg atc agc ggc acc aac gca cac gtc atc ctc gaa caa Ser Ser Phe Gly Ile Ser Gly Thr Asn Ala His Val Ile Leu Glu Gln 2195 2200 2205			6624
gcc tcg aaa aca cca gac gag act gcg gac aag agc ggt ccc gat tcg Ala Ser Lys Thr Pro Asp Glu Thr Ala Asp Lys Ser Gly Pro Asp Ser			6672

2210	2215	2220	
gaa tcg acc gtg gac ctt cca gcg gtc ccg ttg atc gtg tcg ggg aga			6720
Glu Ser Thr Val Asp Leu Pro Ala Val Pro Leu Ile Val Ser Gly Arg			
2225	2230	2235	2240
aca ccg gca gcg ctc agc gct cag gcg agc gca ttg ttg tcc tat ttg			6768
Thr Pro Ala Ala Leu Ser Ala Gln Ala Ser Ala Leu Leu Ser Tyr Leu			
2245	2250		2255
ggt gag cgt ggc gat att tcc acg ctg gat gcg gcg ttt tcg ttg gct			6816
Gly Glu Arg Gly Asp Ile Ser Thr Leu Asp Ala Ala Phe Ser Leu Ala			
2260	2265		2270
tcc tcc cgg gcc gcg ttg gag gag cgg gcg gtg gtg ctg gga gcg gac			6864
Ser Ser Arg Ala Ala Leu Glu Glu Arg Ala Val Val Leu Gly Ala Asp			
2275	2280		2285
cgc gaa acg ttg ttg tcc ggg ttg gaa gcg ctg gct tcc ggt cgc gag			6912
Arg Glu Thr Leu Leu Ser Gly Leu Glu Ala Leu Ala Ser Gly Arg Glu			
2290	2295		2300
gct tct ggg gtg gtg tcg gga tcc ccg gtc tct ggc ggg gtt ggg ttc			6960
Ala Ser Gly Val Val Ser Gly Ser Pro Val Ser Gly Gly Val Gly Phe			
2305	2310	2315	2320
gtg ttc gcc ggt cag ggc gga cag tgg ttg ggg atg ggc ccg ggg ctc			7008
Val Phe Ala Gly Gln Gly Gly Gln Trp Leu Gly Met Gly Arg Gly Leu			
2325	2330		2335
tac tcg gtt ttt ccg gtg ttc gct gac gcg ttt gac gaa gca tgt gcc			7056
Tyr Ser Val Phe Pro Val Phe Ala Asp Ala Phe Asp Glu Ala Cys Ala			
2340	2345		2350
gga ctg gac gcg cat ctg ggg cag gac gtg ggg gtc ccg gat gtg gtg			7104
Gly Leu Asp Ala His Leu Gly Gln Asp Val Gly Val Arg Asp Val Val			
2355	2360		2365
ttt ggt tcc gac ggg tcc ttg ttg gat ccg acg ctg tgg gcc cag tcg			7152
Phe Gly Ser Asp Gly Ser Leu Leu Asp Arg Thr Leu Trp Ala Gln Ser			
2370	2375		2380
ggt ttg ttc gcg ttg cag gtt ggt ttg ctg agc ctg ctg ggt tcg tgg			7200
Gly Leu Phe Ala Leu Gln Val Gly Leu Leu Ser Leu Leu Gly Ser Trp			
2385	2390	2395	2400
ggt gtc ccg ccg ggt gtg gtg ctg ggc cat tcg gtc ggc gag ttc gcg			7248
Gly Val Arg Pro Gly Val Val Leu Gly His Ser Val Gly Glu Phe Ala			

2405	2410	2415	
gcg gcg gtt gcg gcg gga gtg ttg tcg ttg ccg gat gcg gct cgg atg			7296
Ala Ala Val Ala Ala Gly Val Leu Ser Leu Pro Asp Ala Ala Arg Met			
2420	2425	2430	
gtg gcg ggt cgt gcc cgg ttg atg cag gcg ttg cct tct ggc ggt gcc			7344
Val Ala Gly Arg Ala Arg Leu Met Gln Ala Leu Pro Ser Gly Gly Ala			
2435	2440	2445	
atg ttg gcg gtg gct gct ggt gag gag cag ctg cgg ccg ttg ttg gcc			7392
Met Leu Ala Val Ala Ala Gly Glu Glu Gln Leu Arg Pro Leu Leu Ala			
2450	2455	2460	
gat cgg gtt gat ggt gcg ggt atc gcc gcg gtc aac gct cct gag tcg			7440
Asp Arg Val Asp Gly Ala Gly Ile Ala Ala Val Asn Ala Pro Glu Ser			
2465	2470	2475	2480
gtg gtg ctc tcc ggc gat cgg gag gtg ctt gac gac atc gcc ggc gcg			7488
Val Val Leu Ser Gly Asp Arg Glu Val Leu Asp Asp Ile Ala Gly Ala			
2485	2490	2495	
ctg gat ggg caa ggg att cgg tgg cgg cgg ttg cgg gtt tcg cat gcg			7536
Leu Asp Gly Gln Gly Ile Arg Trp Arg Arg Leu Arg Val Ser His Ala			
2500	2505	2510	
ttt cat tcg tat cgg atg gac ccg atg ttg cag gag ttc gcc gaa atc			7584
Phe His Ser Tyr Arg Met Asp Pro Met Leu Gln Glu Phe Ala Glu Ile			
2515	2520	2525	
gca cgc agc gtg gac tac cgg cgt ggc gac cta ccg gtc gtg tcg acg			7632
Ala Arg Ser Val Asp Tyr Arg Arg Gly Asp Leu Pro Val Val Ser Thr			
2530	2535	2540	
ttg acg ggt gag ctc gac acc gca ggt gtg atg gct acg ccg gag tat			7680
Leu Thr Gly Glu Leu Asp Thr Ala Gly Val Met Ala Thr Pro Glu Tyr			
2545	2550	2555	2560
tgg gtg cgt cag gtt cga gag ccc gtc cgc ttc gcc gac ggc gtc cgg			7728
Trp Val Arg Gln Val Arg Glu Pro Val Arg Phe Ala Asp Gly Val Arg			
2565	2570	2575	
gtg ctc gcg cag caa ggg gtc gcc acg atc ttc gaa ctc ggc cct gat			7776
Val Leu Ala Gln Gln Gly Val Ala Thr Ile Phe Glu Leu Gly Pro Asp			
2580	2585	2590	
gcg acg ctg tcg gcc ctg att ccc gat tgt cat tcg tgg gct gat cag			7824
Ala Thr Leu Ser Ala Leu Ile Pro Asp Cys His Ser Trp Ala Asp Gln			

2595

2600

2605

gcc atg ccg att ccg atg ctg cgt aaa gac cgt acg gaa acc gaa act 7872  
 Ala Met Pro Ile Pro Met Leu Arg Lys Asp Arg Thr Glu Thr Glu Thr  
 2610 2615 2620

gtg gtc gcc gcg gtg gcg cgg gcg cac acg cgt ggt gtt ccg gtc gaa 7920  
 Val Val Ala Ala Val Ala Arg Ala His Thr Arg Gly Val Pro Val Glu  
 2625 2630 2635 2640

tgg tcg gcg tat ttc gcc ggc acc ggg gca cgg cgg gtc gag ttg ccg 7968  
 Trp Ser Ala Tyr Phe Ala Gly Thr Gly Ala Arg Arg Val Glu Leu Pro  
 2645 2650 2655

acg tat gcc ttc cag cgg cag cgg tac tgg ctg gaa aca tcg gat tac 8016  
 Thr Tyr Ala Phe Gln Arg Gln Arg Tyr Trp Leu Glu Thr Ser Asp Tyr  
 2660 2665 2670

ggc gat gtg acg ggt atc ggc ctg gct gcg gcg gag cat ccg ttg ctg 8064  
 Gly Asp Val Thr Gly Ile Gly Leu Ala Ala Ala Glu His Pro Leu Leu  
 2675 2680 2685

ggg gcc gtg gtt gcg ctg gcc gat ggt gat ggg atg gtg ctg acc ggc 8112  
 Gly Ala Val Val Ala Leu Ala Asp Gly Asp Gly Met Val Leu Thr Gly  
 2690 2695 2700

cgg ttg tcg gtg ggg acg cat ccg tgg ctg gcc cag cat cgc gtg ctg 8160  
 Arg Leu Ser Val Gly Thr His Pro Trp Leu Ala Gln His Arg Val Leu  
 2705 2710 2715 2720

ggc gag gtc gtc gtc ccc ggc acc gcc atc ctg gag atg gcc ctg cac 8208  
 Gly Glu Val Val Val Pro Gly Thr Ala Ile Leu Glu Met Ala Leu His  
 2725 2730 2735

gca ggg gcg cgt ctc ggc tgt gac cgg gtg gaa gag ctc acc ctg gaa 8256  
 Ala Gly Ala Arg Leu Gly Cys Asp Arg Val Glu Glu Leu Thr Leu Glu  
 2740 2745 2750

aca ccg ctg gtg gtc ccc gaa cgc gcg gcg ggt gcc ggt agt cgt ggc 8304  
 Thr Pro Leu Val Val Pro Glu Arg Ala Ala Gly Ala Gly Ser Arg Gly  
 2755 2760 2765

cct gcg gga ggg acc aca gtt tca att gaa act gcg gaa gaa cgt gtg 8352  
 Pro Ala Gly Gly Thr Thr Val Ser Ile Glu Thr Ala Glu Glu Arg Val  
 2770 2775 2780

cgg acg aac gac gcc atc gaa atc cag ctg ctg gtg aac gca ccc gac 8400  
 Arg Thr Asn Asp Ala Ile Glu Ile Gln Leu Leu Val Asn Ala Pro Asp



2785	2790	2795	2800	
gaa ggc ggt cgg cga agg gtg tcg ctg tat tcc cgc ccg gcc ggt ggg				8448
Glu Gly Gly Arg Arg Arg Val Ser Leu Tyr Ser Arg Pro Ala Gly Gly				
2805	2810	2815		
tcg aga ggt ggg ggt tgg acg cgc cac gcc acc ggc gaa ctc gtc gtc				8496
Ser Arg Gly Gly Gly Trp Thr Arg His Ala Thr Gly Glu Leu Val Val				
2820	2825	2830		
ggc acc acc ggt ggt agg gcg gtt cct gat tgg tcg gct gag ggt gcc				8544
Gly Thr Thr Gly Gly Arg Ala Val Pro Asp Trp Ser Ala Glu Gly Ala				
2835	2840	2845		
gag tcg att gct ctc gat gag ttc tac gtc gct ctg gcc gga aac ggg				8592
Glu Ser Ile Ala Leu Asp Glu Phe Tyr Val Ala Leu Ala Gly Asn Gly				
2850	2855	2860		
ttc gag tac ggg ccg ttg ttc cag ggg ctt cag gcg gca tgg cgt cgt				8640
Phe Glu Tyr Gly Pro Leu Phe Gln Gly Leu Gln Ala Ala Trp Arg Arg				
2865	2870	2875	2880	
ggt gac gag gtt ctc gcc gaa atc gcc ccg ccg gcc gag gcc gat gcg				8688
Gly Asp Glu Val Leu Ala Glu Ile Ala Pro Pro Ala Glu Ala Asp Ala				
2885	2890	2895		
atg gcg tcg gga tac ctg ctc gac cca gcg ttg ctg gat gcc gcg ctg				8736
Met Ala Ser Gly Tyr Leu Leu Asp Pro Ala Leu Leu Asp Ala Ala Leu				
2900	2905	2910		
cag gcg tcc gcg ctc ggc gac cgc ccg gag caa ggc ggc gcg tgg ctg				8784
Gln Ala Ser Ala Leu Gly Asp Arg Pro Glu Gln Gly Gly Ala Trp Leu				
2915	2920	2925		
ccg ttc tca ttc acc ggc gtc gaa ctt tcc gct ccg gca ggg acg atc				8832
Pro Phe Ser Phe Thr Gly Val Glu Leu Ser Ala Pro Ala Gly Thr Ile				
2930	2935	2940		
agc agg gtg cgg ctg gag acc agg cga ccc gac gcg ata tcg gtg gcc				8880
Ser Arg Val Arg Leu Glu Thr Arg Arg Pro Asp Ala Ile Ser Val Ala				
2945	2950	2955	2960	
gtg atg gat gag agt ggg cgg ttg ctc gcc tcg atc gat tct ctc agg				8928
Val Met Asp Glu Ser Gly Arg Leu Leu Ala Ser Ile Asp Ser Leu Arg				
2965	2970	2975		
cta cga agc gtg tcg tcg gga cag ctg gcg aat cgg gac gct gtc cgc				8976
Leu Arg Ser Val Ser Ser Gly Gln Leu Ala Asn Arg Asp Ala Val Arg				

2980

2985

2990

gac gcg ctg ttc gag gtg acc tgg gag ccg gtg gcg acg cag tcg acg 9024  
 Asp Ala Leu Phe Glu Val Thr Trp Glu Pro Val Ala Thr Gln Ser Thr  
 2995 3000 3005

gaa ccg ggt cgc tgg gcc ctg ctt ggt gat act gcc tgc ggt aaa gac 9072  
 Glu Pro Gly Arg Trp Ala Leu Leu Gly Asp Thr Ala Cys Gly Lys Asp  
 3010 3015 3020

gat ctc atc aaa ctc gca acg gat tcc gcc gac cgc tgc gcg gat ctg 9120  
 Asp Leu Ile Lys Leu Ala Thr Asp Ser Ala Asp Arg Cys Ala Asp Leu  
 3025 3030 3035 3040

gcg gcg cta gcc gag aaa ctt gat tcc agc gcg ctg gtt cct gat gtc 9168  
 Ala Ala Leu Ala Glu Lys Leu Asp Ser Ser Ala Leu Val Pro Asp Val  
 3045 3050 3055

gtg gtc tac tgc gcc gga gaa cag gcg gat ccc ggc acc ggc gca gcc 9216  
 Val Val Tyr Cys Ala Gly Glu Gln Ala Asp Pro Gly Thr Gly Ala Ala  
 3060 3065 3070

gca ctt gcg gag acc cag cag acg ttg gct ctg ctc caa gcg tgg ttg 9264  
 Ala Leu Ala Glu Thr Gln Gln Thr Leu Ala Leu Leu Gln Ala Trp Leu  
 3075 3080 3085

gct gag ccg cgg ttg gcc gag gca cgt ctg gtg gtg gtg acg tgt gca 9312  
 Ala Glu Pro Arg Leu Ala Glu Ala Arg Leu Val Val Val Thr Cys Ala  
 3090 3095 3100

gcg gtg acg acg gct ccg agt gac ggt gca tca gag ctg gca cat gcg 9360  
 Ala Val Thr Thr Ala Pro Ser Asp Gly Ala Ser Glu Leu Ala His Ala  
 3105 3110 3115 3120

ccg ttg tgg ggg ttg ttg cgt gcc gcg cag gtg gag aac ccg ggg cag 9408  
 Pro Leu Trp Gly Leu Leu Arg Ala Ala Gln Val Glu Asn Pro Gly Gln  
 3125 3130 3135

ttt gtg ctg gcg gac gtc gac gga acc gcc gaa tcg tgg cgt gcg ttg 9456  
 Phe Val Leu Ala Asp Val Asp Gly Thr Ala Glu Ser Trp Arg Ala Leu  
 3140 3145 3150

ccg agt gcg ttg ggc tcg atg gaa ccg cag ttg gcc ctg ccg aag gcc 9504  
 Pro Ser Ala Leu Gly Ser Met Glu Pro Gln Leu Ala Leu Arg Lys Gly  
 3155 3160 3165

gcg gtg cga gcg ccc cgc ttg gct tcg gtc gcc ggg cag atc gac gtg 9552  
 Ala Val Arg Ala Pro Arg Leu Ala Ser Val Ala Gly Gln Ile Asp Val

3170	3175	3180	
ccc gcg gtt gtg gcg gat ccc gac cga acc gtg ctg att tcg ggc ggc			9600
Pro Ala Val Val Ala Asp Pro Asp Arg Thr Val Leu Ile Ser Gly Gly			
3185	3190	3195	3200
acg ggc ctg ttg ggg ggc gcg gtt gcc cgc cac ctg gtg acc gaa cgc			9648
Thr Gly Leu Leu Gly Gly Ala Val Ala Arg His Leu Val Thr Glu Arg			
3205	3210	3215	
ggt gtc cgc cga ttg gtg ttg acg ggc cgt cgt ggc tgg gat gct cct			9696
Gly Val Arg Arg Leu Val Leu Thr Gly Arg Arg Gly Trp Asp Ala Pro			
3220	3225	3230	
gga atc acc gag ttg gtg ggt gag ctg aac ggc ctc ggt gcc gtg gtc			9744
Gly Ile Thr Glu Leu Val Gly Glu Leu Asn Gly Leu Gly Ala Val Val			
3235	3240	3245	
gac gtg gtg gcg tgc gac gtc gcg gat cgt gct gat ctg gag tcg ttg			9792
Asp Val Val Ala Cys Asp Val Ala Asp Arg Ala Asp Leu Glu Ser Leu			
3250	3255	3260	
ctg gcg gcg gtc ccg gcg gaa ttt ccg ttg tgc ggc gtg gtg cat gcc			9840
Leu Ala Ala Val Pro Ala Glu Phe Pro Leu Cys Gly Val Val His Ala			
3265	3270	3275	3280
gcg ggg gcg ctg gcc gac ggg gtg atc gag tcg ttg tca ccg gac gac			9888
Ala Gly Ala Leu Ala Asp Gly Val Ile Glu Ser Leu Ser Pro Asp Asp			
3285	3290	3295	
gtg gga gcg gtg ttc ggc ccg aag gcg gcg ggg gcg tgg aat ctg cac			9936
Val Gly Ala Val Phe Gly Pro Lys Ala Ala Gly Ala Trp Asn Leu His			
3300	3305	3310	
gag ctg act cgt gat acg gac ctg tcg ttc ttc gcg ttg ttc tcc tcg			9984
Glu Leu Thr Arg Asp Thr Asp Leu Ser Phe Phe Ala Leu Phe Ser Ser			
3315	3320	3325	
ctt tcc ggt gtt gcc ggc gct cct ggt cag ggc aat tat gcg gcg gcg			10032
Leu Ser Gly Val Ala Gly Ala Pro Gly Gln Gly Asn Tyr Ala Ala Ala			
3330	3335	3340	
aac gcg ttc ctg gac gca ttg gcg cat tac cgg cgg tca cag gga ctg			10080
Asn Ala Phe Leu Asp Ala Leu Ala His Tyr Arg Arg Ser Gln Gly Leu			
3345	3350	3355	3360
cct gcg gtg tcg ctg gcc tgg ggc ctg tgg gag cag ccg agc ggg atg			10128
Pro Ala Val Ser Leu Ala Trp Gly Leu Trp Glu Gln Pro Ser Gly Met			

3365

3370

3375

acg gag acg ctc agc gag gtc gac cgg agc agg atc gcg cgc gcc aac 10176  
Thr Glu Thr Leu Ser Glu Val Asp Arg Ser Arg Ile Ala Arg Ala Asn  
3380 3385 3390

ccg ccg ttg tcc acc aag gag gga ttg cgg ctg ttc gat gcc ggg ctg 10224  
Pro Pro Leu Ser Thr Lys Glu Gly Leu Arg Leu Phe Asp Ala Gly Leu  
3395 3400 3405

gcg ctg gac cgg gca gcg gta gtt ccg gcg aag ttg gac agg act ttc 10272  
Ala Leu Asp Arg Ala Ala Val Val Pro Ala Lys Leu Asp Arg Thr Phe  
3410 3415 3420

ctg gcc gag cag gcg cgg tcg ggc tcg ctg ccc gca ttg ttg acg gca 10320  
Leu Ala Glu Gln Ala Arg Ser Gly Ser Leu Pro Ala Leu Leu Thr Ala  
3425 3430 3435 3440

ctg gta ccc ccc atc cgt cgt aat agg cgg gct agc gga acc gag ctc 10368  
Leu Val Pro Pro Ile Arg Arg Asn Arg Arg Ala Ser Gly Thr Glu Leu  
3445 3450 3455

gcg gac gag ggc acc ctg ctc ggg gtg gtg cgg gag cat gcc gcg gcc 10416  
Ala Asp Glu Gly Thr Leu Leu Gly Val Val Arg Glu His Ala Ala Ala  
3460 3465 3470

gtg ctg ggg tat tcg agc gcg gct gac gtc ggg gtc gag cgc gct ttc 10464  
Val Leu Gly Tyr Ser Ser Ala Ala Asp Val Gly Val Glu Arg Ala Phe  
3475 3480 3485

cgg gat ctg ggt ttt gat tcg ttg tct ggt gtg gag ttg cgg aac cgc 10512  
Arg Asp Leu Gly Phe Asp Ser Leu Ser Gly Val Glu Leu Arg Asn Arg  
3490 3495 3500

ctt gcc ggg gtg ctg ggg gtg cgg ttg ccg gcg act gcg gtg ttc gac 10560  
Leu Ala Gly Val Leu Gly Val Arg Leu Pro Ala Thr Ala Val Phe Asp  
3505 3510 3515 3520

tat ccg acg ccg agg gcg ctg gcc cgg ttc ctg cac cag gaa ctg gca 10608  
Tyr Pro Thr Pro Arg Ala Leu Ala Arg Phe Leu His Gln Glu Leu Ala  
3525 3530 3535

gac gag atc gct acg acg cca gcg ccg gtg acg acg acc agg gca ccg 10656  
Asp Glu Ile Ala Thr Thr Pro Ala Pro Val Thr Thr Thr Arg Ala Pro  
3540 3545 3550

gtc gcc gaa gac gat ctc gtc gcg ata gtc ggg atg gga tgc cgt ttt 10704  
Val Ala Glu Asp Asp Leu Val Ala Ile Val Gly Met Gly Cys Arg Phe

3555	3560	3565	
ccc ggt cag gtg tcc tcg ccg gag gag ctc tgg cgt ttg gtg gcc ggg			10752
Pro Gly Gln Val Ser Ser Pro Glu Glu Leu Trp Arg Leu Val Ala Gly			
3570	3575	3580	
ggc gtg gat gcg gtc gcg gac ttc cca gcc gat cgc ggc tgg gat ctg			10800
Gly Val Asp Ala Val Ala Asp Phe Pro Ala Asp Arg Gly Trp Asp Leu			
3585	3590	3595	3600
gca ggc ttg ttc gat ccg gac ccg gaa cgg gct ggg aag acc tac gtg			10848
Ala Gly Leu Phe Asp Pro Asp Pro Glu Arg Ala Gly Lys Thr Tyr Val			
3605	3610	3615	
cgg gaa ggg gcc ttc ctc acc gac gcc gat cgg ttc gat gcg ggt ttc			10896
Arg Glu Gly Ala Phe Leu Thr Asp Ala Asp Arg Phe Asp Ala Gly Phe			
3620	3625	3630	
ctc ggg att tcc ccg cgt gag gcg ttg gcg atg gat ccg cag caa cgg			10944
Phe Gly Ile Ser Pro Arg Glu Ala Leu Ala Met Asp Pro Gln Gln Arg			
3635	3640	3645	
ctg ttg ctg gag ctg tcc tgg gag gcc att gaa cgg gca ggg atc gat			10992
Leu Leu Leu Glu Leu Ser Trp Glu Ala Ile Glu Arg Ala Gly Ile Asp			
3650	3655	3660	
ccg ggt tcg ctg agg ggg agt cgg acc ggt gtg ttc gcg ggg ctg atg			11040
Pro Gly Ser Leu Arg Gly Ser Arg Thr Gly Val Phe Ala Gly Leu Met			
3665	3670	3675	3680
tac cac gac tat ggc gcc cgg ttc gcc agc cga gcc ccg gaa ggt ttc			11088
Tyr His Asp Tyr Gly Ala Arg Phe Ala Ser Arg Ala Pro Glu Gly Phe			
3685	3690	3695	
gag ggg tat ctc ggc aat ggc agt gct ggg agt gtc gcg tcg ggc cgg			11136
Glu Gly Tyr Leu Gly Asn Gly Ser Ala Gly Ser Val Ala Ser Gly Arg			
3700	3705	3710	
att gcg tac tcg ttt ggt ttc gag ggt cct gcg gtg acg gtg gat act			11184
Ile Ala Tyr Ser Phe Gly Phe Glu Gly Pro Ala Val Thr Val Asp Thr			
3715	3720	3725	
gcg tgt tcg tcg tcg ttg gtg gcg ttg cat ttg gcg ggt cag tcg ttg			11232
Ala Cys Ser Ser Ser Leu Val Ala Leu His Leu Ala Gly Gln Ser Leu			
3730	3735	3740	
cgt tcc ggc gaa tgc gat ctc gcc ctt gcc ggt ggt gtg acg gtg atg			11280
Arg Ser Gly Glu Cys Asp Leu Ala Leu Ala Gly Gly Val Thr Val Met			

3745

3750

3755

3760

tcg acg ccc ggg acg ttt gtg gaa ttc tcc cgt cag cgg ggc ctg gca 11328  
 Ser Thr Pro Gly Thr Phe Val Glu Phe Ser Arg Gln Arg Gly Leu Ala  
 3765 3770 3775

ccg gac ggg cgg tgc aag tcg ttc gcg gag agc gcg gac ggt acc ggt 11376  
 Pro Asp Gly Arg Cys Lys Ser Phe Ala Glu Ser Ala Asp Gly Thr Gly  
 3780 3785 3790

tgg ggt gag ggt gct ggt ttg gtg ttg ttg gag cgg ttg tcg gat gct 11424  
 Trp Gly Glu Gly Ala Gly Leu Val Leu Leu Glu Arg Leu Ser Asp Ala  
 3795 3800 3805

cgg cgg aat ggg cat cgg gtg ttg gcg gtg gtt cgt ggg tcg gcg gtg 11472  
 Arg Arg Asn Gly His Arg Val Leu Ala Val Val Arg Gly Ser Ala Val  
 3810 3815 3820

aat cag gat ggt gcg tcg aat ggc ttg acc gcg ccg aat ggt ccc tcg 11520  
 Asn Gln Asp Gly Ala Ser Asn Gly Leu Thr Ala Pro Asn Gly Pro Ser  
 3825 3830 3835 3840

cag cag cgg gtc atc cag cag gcg ttg gcg agt gcg ggt ctg tcg gtg 11568  
 Gln Gln Arg Val Ile Gln Gln Ala Leu Ala Ser Ala Gly Leu Ser Val  
 3845 3850 3855

tcc gat gtg gat gcc gtg gag gcg cat ggg acc ggg acc agg ttg ggt 11616  
 Ser Asp Val Asp Ala Val Glu Ala His Gly Thr Gly Thr Arg Leu Gly  
 3860 3865 3870

gat ccg att gag gcg cag gct ctg att gct acg tat ggg cgc gat cgt 11664  
 Asp Pro Ile Glu Ala Gln Ala Leu Ile Ala Thr Tyr Gly Arg Asp Arg  
 3875 3880 3885

gat ccc ggt cgg ccg ttg tgg ttg ggg tcg gtg aag tcc aac atc ggt 11712  
 Asp Pro Gly Arg Pro Leu Trp Leu Gly Ser Val Lys Ser Asn Ile Gly  
 3890 3895 3900

cat acg cag gcg gcg gcg ggt gtt gcc ggt gtg atc aag atg gtg atg 11760  
 His Thr Gln Ala Ala Ala Gly Val Ala Gly Val Ile Lys Met Val Met  
 3905 3910 3915 3920

gcg atg cgg cac ggg caa ctt ccg cgc acg ctg cac gtg gat gca ccc 11808  
 Ala Met Arg His Gly Gln Leu Pro Arg Thr Leu His Val Asp Ala Pro  
 3925 3930 3935

tcc tcg cag gtg gat tgg tcg gcg ggg agg gtc cag ctc ctg acg gag 11856  
 Ser Ser Gln Val Asp Trp Ser Ala Gly Arg Val Gln Leu Leu Thr Glu

3940	3945	3950	
aac acg ccc tgg ccc gac agt ggt cgc ccc tgt cgg gtg ggg gtg tcg			11904
Asn Thr Pro Trp Pro Asp Ser Gly Arg Pro Cys Arg Val Gly Val Ser			
3955	3960	3965	
tcg ttc ggg atc agc ggc acc aac gcg cac gtc atc ctg gaa cag tcc			11952
Ser Phe Gly Ile Ser Gly Thr Asn Ala His Val Ile Leu Glu Gln Ser			
3970	3975	3980	
acg ggg cag atg gat cag gca gcg gag ccg gat tcg agt cct gtt ctg			12000
Thr Gly Gln Met Asp Gln Ala Ala Glu Pro Asp Ser Ser Pro Val Leu			
3985	3990	3995	4000
gat gtt ccg gtg gtg ccg tgg gtg gtg tcg ggc aaa aca ccc gaa gcg			12048
Asp Val Pro Val Val Pro Trp Val Val Ser Gly Lys Thr Pro Glu Ala			
4005	4010	4015	
cta tcc gcc cag gcg gca acg ttg gcg acc tat ttg gac caa aat gtt			12096
Leu Ser Ala Gln Ala Ala Thr Leu Ala Thr Tyr Leu Asp Gln Asn Val			
4020	4025	4030	
gat gtc tcc cct ctg gac gtt ggg att tcg ctt gcg gtg acc cgt tcg			12144
Asp Val Ser Pro Leu Asp Val Gly Ile Ser Leu Ala Val Thr Arg Ser			
4035	4040	4045	
gcg ctg gat gag ccg gcg gtg gtg ctg ggg tcg gat cgt gac acg ttg			12192
Ala Leu Asp Glu Arg Ala Val Val Leu Gly Ser Asp Arg Asp Thr Leu			
4050	4055	4060	
ttg tct ggc ctg aat gcg ctg gct gcc ggt cat gag gct gct ggc gtg			12240
Leu Ser Gly Leu Asn Ala Leu Ala Ala Gly His Glu Ala Ala Gly Val			
4065	4070	4075	4080
gtt acg gga cct gtc ggg att ggt ggc ccg acc ggg ttt gtg ttc gcc			12288
Val Thr Gly Pro Val Gly Ile Gly Gly Arg Thr Gly Phe Val Phe Ala			
4085	4090	4095	
ggt caa ggc ggt cag tgg ttg ggg atg ggc cgc ccg ttg tac tcg gag			12336
Gly Gln Gly Gly Gln Trp Leu Gly Met Gly Arg Arg Leu Tyr Ser Glu			
4100	4105	4110	
ttt ccg gcg ttc gcc ggt gct ttc gac gaa gca tgc gcc gag ctc gat			12384
Phe Pro Ala Phe Ala Gly Ala Phe Asp Glu Ala Cys Ala Glu Leu Asp			
4115	4120	4125	
gcg aac ctg ggg agg gaa gtc ggg gtt ccg gat gtg gtg ttc ggc tcc			12432
Ala Asn Leu Gly Arg Glu Val Gly Val Arg Asp Val Val Phe Gly Ser			

4130	4135	4140	
gac gag tcc ttg ctg gat cgg act ttg tgg gcg cag tcg ggt ttg ttc			12480
Asp Glu Ser Leu Leu Asp Arg Thr Leu Trp Ala Gln Ser Gly Leu Phe			
4145	4150	4155	4160
gcg ttg cag gtc ggt ctc tgg gaa ttg ttg ggt acg tgg ggt gtt cgg			12528
Ala Leu Gln Val Gly Leu Trp Glu Leu Leu Gly Thr Trp Gly Val Arg			
4165	4170		4175
ccc agc gta gtg ctg ggg cat tcg gtc ggg gag cta gcc gcg gcg ttc			12576
Pro Ser Val Val Leu Gly His Ser Val Gly Glu Leu Ala Ala Ala Phe			
4180	4185		4190
gcc gca ggt gtg ctg tcg atg gcg gag gcg gct cgg ctg gtg gcg ggt			12624
Ala Ala Gly Val Leu Ser Met Ala Glu Ala Ala Arg Leu Val Ala Gly			
4195	4200		4205
cgt gcg cgg ttg atg cag gcg ttg cct tct ggc ggt gcc atg ctg gcg			12672
Arg Ala Arg Leu Met Gln Ala Leu Pro Ser Gly Gly Ala Met Leu Ala			
4210	4215		4220
gtg tcc gcg acc gag gcc cga gtc ggc ccg ctg ctc gat ggg gtg cgg			12720
Val Ser Ala Thr Glu Ala Arg Val Gly Pro Leu Leu Asp Gly Val Arg			
4225	4230	4235	4240
gat cgt gtt ggt gtc gca gcg gtt aac gct ccg ggg tcg gtg gtg ctt			12768
Asp Arg Val Gly Val Ala Ala Val Asn Ala Pro Gly Ser Val Val Leu			
4245	4250		4255
tcc ggt gac cgg gat gtg ctc gat ggc att gcc ggt cgg ctg gac ggg			12816
Ser Gly Asp Arg Asp Val Leu Asp Gly Ile Ala Gly Arg Leu Asp Gly			
4260	4265		4270
caa ggt atc cgg tcg agg tgg ttg cgg gtt tcg cac gcg ttt cat tcg			12864
Gln Gly Ile Arg Ser Arg Trp Leu Arg Val Ser His Ala Phe His Ser			
4275	4280		4285
cat cgg atg gat ccg atg ctg gcg gag ttc gcc gag ctc gca cgg agc			12912
His Arg Met Asp Pro Met Leu Ala Glu Phe Ala Glu Leu Ala Arg Ser			
4290	4295	4300	
gtg gac tac cgg tct cca cgg ctg ccg att gtc tcg acg ctg acc gga			12960
Val Asp Tyr Arg Ser Pro Arg Leu Pro Ile Val Ser Thr Leu Thr Gly			
4305	4310	4315	4320
aac ctc gat gac gtg ggc gtg atg gct acg ccg gag tat tgg gtg cgc			13008
Asn Leu Asp Asp Val Gly Val Met Ala Thr Pro Glu Tyr Trp Val Arg			



4325	4330	4335	
cag gtg cga gag ccc gtc cgc ttc gcc gac ggt gtc cag gcg ctt gtg			13056
Gln Val Arg Glu Pro Val Arg Phe Ala Asp Gly Val Gln Ala Leu Val			
4340	4345	4350	
gac caa ggc gtc gac acg att gtg gaa ctc ggt ccg gac ggg gcg ttg			13104
Asp Gln Gly Val Asp Thr Ile Val Glu Leu Gly Pro Asp Gly Ala Leu			
4355	4360	4365	
tcg agc ttg gtt caa gag tgt gtg gcg gag tcc ggg cgg gcg acg ggg			13152
Ser Ser Leu Val Gln Glu Cys Val Ala Glu Ser Gly Arg Ala Thr Gly			
4370	4375	4380	
att ccg ttg gtg cgg aga gac cgt gat gag gtc cga acg gtg ctg gac			13200
Ile Pro Leu Val Arg Arg Asp Arg Asp Glu Val Arg Thr Val Leu Asp			
4385	4390	4395	4400
gct ttg gcg cag acc cac act cgt ggt ggc gcg gtg gac tgg ggg tca			13248
Ala Leu Ala Gln Thr His Thr Arg Gly Gly Ala Val Asp Trp Gly Ser			
4405	4410	4415	
ttt ttc gct ggt acg agg gca acg caa gtc gac ctt ccc acg tat gcc			13296
Phe Phe Ala Gly Thr Arg Ala Thr Gln Val Asp Leu Pro Thr Tyr Ala			
4420	4425	4430	
ttc caa cga cag cgg tac tgg ctg gag cca tcg gat tcc ggt gat gtg			13344
Phe Gln Arg Gln Arg Tyr Trp Leu Glu Pro Ser Asp Ser Gly Asp Val			
4435	4440	4445	
acc ggt gtt ggc ctg acc ggg gcg gag cat ccg ctg ttg ggt gcc gtg			13392
Thr Gly Val Gly Leu Thr Gly Ala Glu His Pro Leu Leu Gly Ala Val			
4450	4455	4460	
gtg ccg gtc gcg ggc ggc gat gag gtg ctg ctg acc ggc agg ctg tcg			13440
Val Pro Val Ala Gly Gly Asp Glu Val Leu Leu Thr Gly Arg Leu Ser			
4465	4470	4475	4480
gtg ggg acg cat ccg tgg ctg gcg gaa cac cgc gtg ctg ggc gaa gtc			13488
Val Gly Thr His Pro Trp Leu Ala Glu His Arg Val Leu Gly Glu Val			
4485	4490	4495	
gtc gtc ccc ggc acc gcg ttg ctg gag atg gcg tgg cgg gcc ggt agc			13536
Val Val Pro Gly Thr Ala Leu Leu Glu Met Ala Trp Arg Ala Gly Ser			
4500	4505	4510	
cag gtc ggt tgt gaa cgt gtg gag gag ctc acc ttg gag gca ccg ctg			13584
Gln Val Gly Cys Glu Arg Val Glu Glu Leu Thr Leu Glu Ala Pro Leu			

4515

4520

4525

gtc ctg ccg gag cgg ggc gct gcg gcg gtg cag ttg gcg gtg ggg gct 13632  
 Val Leu Pro Glu Arg Gly Ala Ala Val Gln Leu Ala Val Gly Ala  
 4530 4535 4540

ccg gat gag gcc ggc cgg cgc agt ttg cag ctc tat tcc cga ggc gct 13680  
 Pro Asp Glu Ala Gly Arg Arg Ser Leu Gln Leu Tyr Ser Arg Gly Ala  
 4545 4550 4555 4560

gat gaa gac ggc gac tgg cgg cgg att gcc tcc ggg ctg ttg gcc cag 13728  
 Asp Glu Asp Gly Asp Trp Arg Arg Ile Ala Ser Gly Leu Leu Ala Gln  
 4565 4570 4575

gcc aat gcg gtg ccg ccg gcg gat tcg acg gca tgg ccg ccg gac ggc 13776  
 Ala Asn Ala Val Pro Pro Ala Asp Ser Thr Ala Trp Pro Pro Asp Gly  
 4580 4585 4590

gcc ggg cag gtc gat ctg gcg gag ttc tac gag cgc ctc gcc gag cgc 13824  
 Ala Gly Gln Val Asp Leu Ala Glu Phe Tyr Glu Arg Leu Ala Glu Arg  
 4595 4600 4605

ggc ttg acc tac ggt ccg gta ttc caa ggg ctc cgc gcc gca tgg cgg 13872  
 Gly Leu Thr Tyr Gly Pro Val Phe Gln Gly Leu Arg Ala Ala Trp Arg  
 4610 4615 4620

cac ggc gac gat atc ttc gcc gaa ttg gcc ggg tca cca gac gcc tcg 13920  
 His Gly Asp Asp Ile Phe Ala Glu Leu Ala Gly Ser Pro Asp Ala Ser  
 4625 4630 4635 4640

ggt ttc ggc atc cac ccg gcg ctg ctg gac gct gca ctg cac gcg atg 13968  
 Gly Phe Gly Ile His Pro Ala Leu Leu Asp Ala Ala Leu His Ala Met  
 4645 4650 4655

gcg ctt ggt gct tcg ccc gac tcg gaa gcg cgt ctg ccg ttt tcc tgg 14016  
 Ala Leu Gly Ala Ser Pro Asp Ser Glu Ala Arg Leu Pro Phe Ser Trp  
 4660 4665 4670

cgt ggc gcc cag ctg tac cgc gct gaa gga gca gcg ctt cgg gta cgg 14064  
 Arg Gly Ala Gln Leu Tyr Arg Ala Glu Gly Ala Ala Leu Arg Val Arg  
 4675 4680 4685

ctc tcg ccg ctg ggc tcc ggt gca gtc tca ttg acg ttg gtg gat gcc 14112  
 Leu Ser Pro Leu Gly Ser Gly Ala Val Ser Leu Thr Leu Val Asp Ala  
 4690 4695 4700

aca ggg cga cga gtc gct gcg gtg gaa tcg ctt tcg acg cga ccg gtc 14160  
 Thr Gly Arg Arg Val Ala Ala Val Glu Ser Leu Ser Thr Arg Pro Val

4705	4710	4715	4720	
tcc acc gac cag atc ggt gcc ggt cgc ggc gat caa gag cgg ctg ctg				14208
Ser Thr Asp Gln Ile Gly Ala Gly Arg Gly Asp Gln Glu Arg Leu Leu				
4725		4730	4735	
cac gtc gag tgg gta agg tcg gct gaa tct gcg ggg atg tct ctg acc				14256
His Val Glu Trp Val Arg Ser Ala Glu Ser Ala Gly Met Ser Leu Thr				
4740	4745		4750	
tcc tgc gcg gtg gtc ggt ttg ggc gaa ccg gag tgg cac gct gcg ctg				14304
Ser Cys Ala Val Val Gly Leu Gly Glu Pro Glu Trp His Ala Ala Leu				
4755	4760		4765	
aag acc act ggt gtc caa gtc gag tcc cat gcg gac ctt gct tcg ttg				14352
Lys Thr Thr Gly Val Gln Val Glu Ser His Ala Asp Leu Ala Ser Leu				
4770	4775		4780	
gcc acc gag gtt gcc aag cgg ggt tca gct cct ggt gcg gtc atc gtc				14400
Ala Thr Glu Val Ala Lys Arg Gly Ser Ala Pro Gly Ala Val Ile Val				
4785	4790	4795	4800	
ccg tgc ccg cga ccc cga gcg atg cag gag ctg ccg acc gcc gcg cga				14448
Pro Cys Pro Arg Pro Arg Ala Met Gln Glu Leu Pro Thr Ala Ala Arg				
4805	4810		4815	
agg gcg acg caa cag gcg atg gcg atg ctg cag caa tgg ctt gcc gat				14496
Arg Ala Thr Gln Gln Ala Met Ala Met Leu Gln Gln Trp Leu Ala Asp				
4820	4825		4830	
gac cgg ttc gtc agt acg cgc ctg atc ctg ctg acg cat cgg gcg gtc				14544
Asp Arg Phe Val Ser Thr Arg Leu Ile Leu Leu Thr His Arg Ala Val				
4835	4840		4845	
tcc gca gtt gct gga gaa gac gtg ctc gac ctg gta cac gcg ccg ctg				14592
Ser Ala Val Ala Gly Glu Asp Val Leu Asp Leu Val His Ala Pro Leu				
4850	4855		4860	
tgg ggc ttg gtc cgc agc gcg caa gcg gag cac ccg gac cga ttc gcc				14640
Trp Gly Leu Val Arg Ser Ala Gln Ala Glu His Pro Asp Arg Phe Ala				
4865	4870	4875	4880	
ttg atc gat atg gac gac gag cga gca tcg cag acg gca ctc gcc gaa				14688
Leu Ile Asp Met Asp Asp Glu Arg Ala Ser Gln Thr Ala Leu Ala Glu				
4885	4890		4895	
gcg ctg act gcg gga gaa gcg cag ctc gcg gtg ccg tcg gga gtt gtg				14736
Ala Leu Thr Ala Gly Glu Ala Gln Leu Ala Val Arg Ser Gly Val Val				

4900	4905	4910	
ctg gcg ccc cgc ctc ggc cag gtg aag gtg agt gga ggt gaa gcg ttc			14784
Leu Ala Pro Arg Leu Gly Gln Val Lys Val Ser Gly Gly Glu Ala Phe			
4915	4920	4925	
agg tgg gat gaa ggc acc gtg ctg gtc acc ggc gga acc ggc ggg ctc			14832
Arg Trp Asp Glu Gly Thr Val Leu Val Thr Gly Gly Thr Gly Gly Leu			
4930	4935	4940	
ggg gcc ctg ctc gca cgc cat ctg gtc agc gcc cac ggt gtg cgg cac			14880
Gly Ala Leu Leu Ala Arg His Leu Val Ser Ala His Gly Val Arg His			
4945	4950	4955	4960
ctg ttg ctc gca agt cgc cgt ggt ctg gcg gcg ccc gga gcg gat gag			14928
Leu Leu Leu Ala Ser Arg Arg Gly Leu Ala Ala Pro Gly Ala Asp Glu			
4965	4970	4975	
ctg gtg gcc gag ctg gag cag gcc ggc gcc gac gtc gcg gtc gtc gcg			14976
Leu Val Ala Glu Leu Glu Gln Ala Gly Ala Asp Val Ala Val Val Ala			
4980	4985	4990	
tgc gac tcg gca gat cgg gac tcg ctt gcg cgg ctg gtg gcg tcg gtg			15024
Cys Asp Ser Ala Asp Arg Asp Ser Leu Ala Arg Leu Val Ala Ser Val			
4995	5000	5005	
cct gcg gaa aac ccg ttg cgg gtg gtg gtg cac gcc gcc ggt gtg ctg			15072
Pro Ala Glu Asn Pro Leu Arg Val Val Val His Ala Ala Gly Val Leu			
5010	5015	5020	
gat gac ggt gtg ctg atg tcg atg tcg ccg gag cgc ttg gac gcg gtg			15120
Asp Asp Gly Val Leu Met Ser Met Ser Pro Glu Arg Leu Asp Ala Val			
5025	5030	5035	5040
ttg cgg ccc aaa gtg gat gcc gcg tgg tac ctg cac gag ctg act cgg			15168
Leu Arg Pro Lys Val Asp Ala Ala Trp Tyr Leu His Glu Leu Thr Arg			
5045	5050	5055	
gaa ctc ggt ctg tcg gcg ttc gtg ttg ttc tcc tcg gtc gcg ggc ctg			15216
Glu Leu Gly Leu Ser Ala Phe Val Leu Phe Ser Ser Val Ala Gly Leu			
5060	5065	5070	
ttc ggc ggt gcg ggg cag agc aat tac gct gcc ggc aac gct ttc ctg			15264
Phe Gly Gly Ala Gly Gln Ser Asn Tyr Ala Ala Gly Asn Ala Phe Leu			
5075	5080	5085	
gat gcc ttg gcg cat tgc cgg cag gcc cag ggg ctg ccc gcg ctg tcg			15312
Asp Ala Leu Ala His Cys Arg Gln Ala Gln Gly Leu Pro Ala Leu Ser			

5090	5095	5100	
ctg gcc tcc ggg ctg tgg gcg agt atc gat gga atg gcg ggc gac ctc			15360
Leu Ala Ser Gly Leu Trp Ala Ser Ile Asp Gly Met Ala Gly Asp Leu			
5105	5110	5115	5120
gct gcg gca gat gtg gag cgg ctg tcg cgg gca ggc att ggc ccg ctt			15408
Ala Ala Ala Asp Val Glu Arg Leu Ser Arg Ala Gly Ile Gly Pro Leu			
	5125	5130	5135
tcg gca ccg gga ggg ctg gcc ttg ttc gac gct gcc gtt ggc tcg gac			15456
Ser Ala Pro Gly Gly Leu Ala Leu Phe Asp Ala Ala Val Gly Ser Asp			
	5140	5145	5150
gaa ccg ttg ctg gca ccg gtg cga ctg gat gtc gaa gca ctg cgt gtg			15504
Glu Pro Leu Leu Ala Pro Val Arg Leu Asp Val Glu Ala Leu Arg Val			
	5155	5160	5165
cag gcc cga tcc gtg cag acc cgg att ccg gaa atg ctg cat ggc atg			15552
Gln Ala Arg Ser Val Gln Thr Arg Ile Pro Glu Met Leu His Gly Met			
	5170	5175	5180
gca atg ggg cca agc cgc cgc act ccg ttc act tcc agg gtt gag ccg			15600
Ala Met Gly Pro Ser Arg Arg Thr Pro Phe Thr Ser Arg Val Glu Pro			
	5185	5190	5195
			5200
ttg cac gaa ccg ctg gcc gga ttg tcg gag ggc gaa cgt ccg cag caa			15648
Leu His Glu Arg Leu Ala Gly Leu Ser Glu Gly Glu Arg Arg Gln Gln			
	5205	5210	5215
gtg ctc cag cgc gtc cgc gcc gat atc gcg gtg gta ctg ggg cac ggc			15696
Val Leu Gln Arg Val Arg Ala Asp Ile Ala Val Val Leu Gly His Gly			
	5220	5225	5230
agg tcg agc gat gtg gac atc gag aag cct ttg gcc gag ctg ggt ttc			15744
Arg Ser Ser Asp Val Asp Ile Glu Lys Pro Leu Ala Glu Leu Gly Phe			
	5235	5240	5245
gac tcg ctg acg gcc atc gaa ctc cgc aac cgt ctc gct acc gcc acc			15792
Asp Ser Leu Thr Ala Ile Glu Leu Arg Asn Arg Leu Ala Thr Ala Thr			
	5250	5255	5260
gga ctg ccg ctt ccc gcg acg ctg gcc ttc gac cac ggc act gcg gcg			15840
Gly Leu Arg Leu Pro Ala Thr Leu Ala Phe Asp His Gly Thr Ala Ala			
	5265	5270	5275
			5280
gca ctc gcc cag cac gtg tgc gcg cag cta ggc acc gcg acc gcg ccg			15888
Ala Leu Ala Gln His Val Cys Ala Gln Leu Gly Thr Ala Thr Ala Pro			

gca ccg agg cga acc gac gac aac gac gcc acg gag ccc gtg agg tcg 15936  
 Ala Pro Arg Arg Thr Asp Asp Asn Asp Ala Thr Glu Pro Val Arg Ser  
 5300 5305 5310

ctc ttc caa cag gcg tat gcg gct ggc cgg ata ctt gac ggg atg gat 15984  
 Leu Phe Gln Gln Ala Tyr Ala Ala Gly Arg Ile Leu Asp Gly Met Asp  
 5315 5320 5325

ttg gtg aag gtc gct gcc cag ttg cga ccg gtg ttc ggt tcg cct ggc 16032  
 Leu Val Lys Val Ala Ala Gln Leu Arg Pro Val Phe Gly Ser Pro Gly  
 5330 5335 5340

gag ctg gaa tcc ctg ccg aaa ccc gtc cag ctt tcc cgt ggt ccc gaa 16080  
 Glu Leu Glu Ser Leu Pro Lys Pro Val Gln Leu Ser Arg Gly Pro Glu  
 5345 5350 5355 5360

gag ctt gcc ttg gtg tgc atg ccg gcg ctg atc ggg atg ccg ccc gca 16128  
 Glu Leu Ala Leu Val Cys Met Pro Ala Leu Ile Gly Met Pro Pro Ala  
 5365 5370 5375

cag cag tac gcg cgg atc gcc gcc ggg ttc cgc gat gtg cgg gac gtt 16176  
 Gln Gln Tyr Ala Arg Ile Ala Ala Gly Phe Arg Asp Val Arg Asp Val  
 5380 5385 5390

tcg gtg atc ccg atg cct gga ttc att gcg gga gaa ccg ctg ccg tcc 16224  
 Ser Val Ile Pro Met Pro Gly Phe Ile Ala Gly Glu Pro Leu Pro Ser  
 5395 5400 5405

gcc atc gag gtg gcg gtt cgg acg cag gcg gag gcg gtg ctg cag gaa 16272  
 Ala Ile Glu Val Ala Val Arg Thr Gln Ala Glu Ala Val Leu Gln Glu  
 5410 5415 5420

ttc gcc ggg ggc tcg ttc gta ctg gtc ggg cat tcc tcc ggg ggc tgg 16320  
 Phe Ala Gly Gly Ser Phe Val Leu Val Gly His Ser Ser Gly Gly Trp  
 5425 5430 5435 5440

ctg gcg cac gag gta gcc ggt gag ctg gag cgt cgc ggg gtc gtc ccg 16368  
 Leu Ala His Glu Val Ala Gly Glu Leu Glu Arg Arg Gly Val Val Pro  
 5445 5450 5455

gcc ggg gtc gta ctg ctg gac acc tac atc ccc ggt gag atc acg ccg 16416  
 Ala Gly Val Val Leu Leu Asp Thr Tyr Ile Pro Gly Glu Ile Thr Pro  
 5460 5465 5470

agg ttc tcc gtg gcg atg gcc cac cgg acg tat gag aag ctc gcg act 16464  
 Arg Phe Ser Val Ala Met Ala His Arg Thr Tyr Glu Lys Leu Ala Thr

5475

5480

5485

ttc acg gac atg cag gat gtc ggt atc acc gcg atg ggc ggg tac ttc 16512  
 Phe Thr Asp Met Gln Asp Val Gly Ile Thr Ala Met Gly Gly Tyr Phe  
 5490 5495 5500

cgg atg ttc acc gag tgg act ccg acg ccg atc ggt gct ccg acg ctg 16560  
 Arg Met Phe Thr Glu Trp Thr Pro Thr Pro Ile Gly Ala Pro Thr Leu  
 5505 5510 5515 5520

ttc gtg cgg acc gaa gat tgc gtc gca gac cct gaa ggg cgg ccg tgg 16608  
 Phe Val Arg Thr Glu Asp Cys Val Ala Asp Pro Glu Gly Arg Pro Trp  
 5525 5530 5535

aca gat gac tcc tgg cgg cca ggg tgg act ctc gcg gat gcc acg gtc 16656  
 Thr Asp Asp Ser Trp Arg Pro Gly Trp Thr Leu Ala Asp Ala Thr Val  
 5540 5545 5550

cag gtg ccg ggc gac cac ttc tcg atg atg gac gag cac gcc ggg tcc 16704  
 Gln Val Pro Gly Asp His Phe Ser Met Met Asp Glu His Ala Gly Ser  
 5555 5560 5565

acc gca cag gca gtc gcg agt tgg ctt gac aaa ctc aac cag cgc acc 16752  
 Thr Ala Gln Ala Val Ala Ser Trp Leu Asp Lys Leu Asn Gln Arg Thr  
 5570 5575 5580

gct cgg caa cgc tga 16767  
 Ala Arg Gln Arg  
 5585

&lt;210&gt; 50

&lt;211&gt; 5588

&lt;212&gt; PRT

&lt;213&gt; Saccharopolyspora spinosa

&lt;400&gt; 50

Met Ala Asn Glu Glu Lys Leu Arg Glu Tyr Leu Lys Arg Val Val Val  
 1 5 10 15

Glu Leu Glu Glu Ala His Glu Arg Leu His Glu Leu Glu Arg Gln Glu  
 20 25 30

His Asp Pro Ile Ala Ile Val Ser Met Gly Cys Arg Tyr Pro Gly Gly  
 35 40 45

Val Ser Thr Pro Glu Glu Leu Trp Arg Leu Val Val Asp Gly Gly Asp  
 50 55 60

Ala Ile Ala Asn Phe Pro Glu Asp Arg Gly Trp Asn Leu Asp Glu Leu  
65 70 75 80

Phe Asp Pro Asp Pro Gly Arg Ala Gly Thr Ser Tyr Val Arg Glu Gly  
85 90 95

Gly Phe Leu Arg Gly Val Ala Asp Phe Asp Ala Gly Leu Phe Gly Ile  
100 105 110

Ser Pro Arg Glu Ala Gln Ala Met Asp Pro Gln Gln Arg Leu Leu Leu  
115 120 125

Glu Ile Ser Trp Glu Val Phe Glu Arg Ala Gly Ile Asp Pro Phe Ser  
130 135 140

Leu Arg Gly Thr Lys Thr Gly Val Phe Ala Gly Leu Ile Tyr His Asp  
145 150 155 160

Tyr Ala Ser Arg Phe Arg Lys Thr Pro Ala Glu Phe Glu Gly Tyr Phe  
165 170 175

Ala Thr Gly Asn Ala Gly Ser Val Ala Ser Gly Arg Val Ala Tyr Thr  
180 185 190

Phe Gly Leu Glu Gly Pro Ala Val Thr Val Asp Thr Ala Cys Ser Ser  
195 200 205

Ser Leu Val Ala Leu His Leu Ala Cys Gln Ser Leu Arg Leu Gly Glu  
210 215 220

Cys Asp Leu Ala Leu Ala Gly Gly Ile Ser Val Met Ala Thr Pro Gly  
225 230 235 240

Ala Phe Val Glu Phe Ser Arg Gln Arg Ala Leu Ala Ser Asp Gly Arg  
245 250 255

Cys Lys Pro Phe Ala Asp Ala Ala Asp Gly Thr Gly Trp Gly Glu Gly  
260 265 270

Ala Gly Met Leu Leu Leu Glu Arg Leu Ser Asp Ala Arg Arg Asn Gly  
275 280 285

His Pro Val Leu Ala Ala Val Val Gly Ser Ala Ile Asn Gln Asp Gly  
290 295 300

Thr Ser Asn Gly Leu Thr Ala Pro Ser Gly Pro Ala Gln Gln Arg Val  
305 310 315 320



Ile Arg Gln Ala Leu Ala Asn Ala Gly Leu Ser Pro Ala Glu Val Asp  
325 330 335

Val Val Glu Ala His Gly Thr Gly Thr Ala Leu Gly Asp Pro Ile Glu  
340 345 350

Ala Gln Ala Leu Ile Ala Thr Tyr Gly Ala Asn Arg Ser Ala Asp His  
355 360 365

Pro Leu Leu Leu Gly Ser Leu Lys Ser Asn Ile Gly His Thr Gln Ala  
370 375 380

Ala Ala Gly Val Ala Gly Val Ile Lys Ser Val Leu Ala Ile Arg His  
385 390 395 400

Arg Glu Met Pro Arg Ser Leu His Ile Asp Gln Pro Ser Gln His Val  
405 410 415

Asp Trp Ser Ala Gly Ala Val Arg Leu Leu Thr Asp Ser Val Asp Trp  
420 425 430

Pro Asp Leu Gly Arg Pro Arg Arg Ala Gly Val Ser Ser Phe Gly Met  
435 440 445

Ser Gly Thr Asn Ala His Leu Ile Val Glu Glu Val Ser Asp Glu Pro  
450 455 460

Val Ser Gly Ser Thr Glu Pro Thr Gly Ala Phe Pro Trp Pro Leu Ser  
465 470 475 480

Gly Lys Thr Glu Thr Ala Leu Arg Glu Gln Ala Ala Glu Leu Leu Ser  
485 490 495

Val Val Thr Glu His Pro Glu Pro Gly Leu Gly Asp Val Gly Tyr Ser  
500 505 510

Leu Ala Thr Gly Arg Ala Ala Met Glu His Arg Ala Val Val Val Ala  
515 520 525

Asp Asp Arg Asp Ser Phe Val Ala Gly Leu Thr Ala Leu Ala Ala Gly  
530 535 540

Val Pro Ala Ala Asn Val Val Gln Gly Ala Ala Asp Cys Lys Gly Lys  
545 550 555 560

Val Ala Phe Val Phe Pro Gly Gln Gly Ser His Trp Gln Gly Met Ala  
565 570 575

Arg Glu Leu Ser Glu Ser Ser Pro Val Phe Arg Arg Lys Leu Ala Glu  
580 585 590

Cys Ala Ala Ala Thr Ala Pro Tyr Val Asp Trp Ser Leu Leu Gly Val  
595 600 605

Leu Arg Gly Asp Pro Asp Ala Pro Ala Leu Asp Arg Asp Asp Val Ile  
610 615 620

Gln Leu Ala Leu Phe Ala Met Met Val Ser Leu Ala Glu Leu Trp Arg  
625 630 635 640

Ser Cys Gly Val Glu Pro Ala Ala Val Val Gly His Ser Gln Gly Glu  
645 650 655

Ile Ala Ala Ala His Val Ala Gly Ala Leu Ser Leu Thr Asp Ala Val  
660 665 670

Arg Ile Ile Ala Ala Arg Cys Asp Ala Val Ser Ala Leu Thr Gly Lys  
675 680 685

Gly Gly Met Leu Ala Ile Ala Leu Pro Glu Ser Ala Val Val Lys Arg  
690 695 700

Ile Ala Gly Leu Pro Glu Leu Thr Val Ala Ala Val Asn Gly Pro Gly  
705 710 715 720

Ser Thr Val Val Ser Gly Glu Pro Ser Ala Leu Glu Arg Leu Gln Thr  
725 730 735

Glu Leu Thr Ala Glu Asn Val Gln Thr Arg Arg Val Gly Ile Asp Tyr  
740 745 750

Ala Ser His Ser Pro Gln Ile Ala Gln Val Gln Gly Arg Leu Leu Asp  
755 760 765

Arg Leu Gly Glu Val Gly Ser Glu Pro Ala Glu Ile Ala Phe Tyr Ser  
770 775 780

Thr Val Thr Gly Glu Arg Thr Asp Thr Gly Arg Leu Asp Ala Asp Tyr  
785 790 795 800

Trp Tyr Gln Asn Leu Arg Gln Pro Val Arg Phe Gln Gln Thr Val Ala  
805 810 815

Arg Met Ala Asp Gln Gly Tyr Arg Phe Phe Val Glu Val Ser Pro His  
820 825 830

Pro Leu Leu Thr Ala Gly Ile Gln Glu Thr Leu Glu Ala Ala Asp Ala  
835 840 845

Gly Gly Val Val Val Gly Ser Leu Arg Arg Gly Glu Gly Gly Ser Arg  
850 855 860

Arg Trp Leu Thr Ser Leu Ala Glu Cys Gln Val Arg Gly Leu Pro Val  
865 870 875 880

Asn Trp Glu Gln Val Phe Leu Asn Thr Gly Ala Arg Arg Val Pro Leu  
885 890 895

Pro Thr Tyr Pro Phe Gln Arg Gln Arg Tyr Trp Leu Glu Ser Ala Glu  
900 905 910

Tyr Asp Ala Gly Asp Leu Gly Ser Val Gly Leu Leu Ser Ala Glu His  
915 920 925

Pro Leu Leu Gly Ala Ala Val Thr Leu Ala Asp Ala Gly Gly Phe Leu  
930 935 940

Leu Thr Gly Lys Leu Ser Val Lys Thr Gln Pro Trp Leu Ala Asp His  
945 950 955 960

Val Val Gly Gly Ala Ile Leu Leu Pro Gly Thr Ala Phe Val Glu Met  
965 970 975

Leu Ile Arg Ala Ala Asp Gln Val Gly Cys Asp Leu Ile Glu Glu Leu  
980 985 990

Ser Leu Thr Thr Pro Leu Val Leu Pro Ala Thr Gly Ala Val Gln Val  
995 1000 1005

Gln Ile Ala Val Gly Gly Pro Asp Glu Ala Gly Arg Arg Ser Val Arg  
1010 1015 1020

Val His Ser Cys Arg Asp Asp Ala Val Pro Gln Asp Ser Trp Thr Cys  
1025 1030 1035 1040

His Ala Thr Gly Thr Leu Thr Ser Ser Asp His Gln Asp Ala Gly Gln  
1045 1050 1055

Gly Pro Asp Gly Ile Trp Pro Pro Asn Asp Ala Val Ala Val Pro Leu  
1060 1065 1070

Asp Ser Phe Tyr Ala Arg Ala Ala Glu Arg Gly Phe Asp Phe Gly Pro  
1075 1080 1085

Ala Phe Gln Gly Leu Gln Ala Ala Trp Lys Arg Gly Asp Glu Ile Phe  
 1090 1095 1100

Ala Glu Val Gly Leu Pro Thr Ala His Arg Glu Asp Ala Gly Arg Phe  
 105 1110 1115 1120

Gly Ile His Pro Ala Leu Leu Asp Ala Ala Leu Gln Ala Leu Gly Ala  
 1125 1130 1135

Ala Glu Glu Asp Pro Asp Glu Gly Trp Leu Pro Phe Ala Trp Gln Gly  
 1140 1145 1150

Val Ser Leu Lys Ala Thr Gly Ala Leu Ser Leu Arg Val His Leu Val  
 1155 1160 1165

Pro Ala Gly Ala Asn Ala Val Ser Val Phe Thr Thr Asp Thr Thr Gly  
 1170 1175 1180

Gln Ala Val Leu Ser Ile Asp Ser Leu Val Leu Arg Gln Ile Ser Asp  
 185 1190 1195 1200

Lys Gln Leu Ala Ala Ala Arg Ala Met Glu His Glu Ser Leu Phe Arg  
 1205 1210 1215

Val Asp Trp Lys Arg Ile Ser Pro Gly Ala Ala Lys Pro Val Ser Trp  
 1220 1225 1230

Ala Val Ile Gly Asn Asp Glu Leu Ala Arg Ala Cys Gly Ser Ala Leu  
 1235 1240 1245

Gly Thr Glu Leu His Pro Asp Leu Thr Gly Leu Ala Asp Pro Pro Pro  
 1250 1255 1260

Asp Val Val Val Val Pro Cys Gly Ala Ser Arg Gln Asp Leu Asp Val  
 265 1270 1275 1280

Ala Ser Glu Ala Arg Ala Ala Thr Gln Arg Met Leu Asp Leu Ile Gln  
 1285 1290 1295

Asp Trp Leu Ala Ala Ala Arg Phe Ala Gly Ser Arg Leu Val Val Val  
 1300 1305 1310

Thr Cys Gly Ala Ala Ser Thr Gly Pro Ala Glu Gly Val Ser Asp Leu  
 1315 1320 1325

Val His Ala Ala Ser Trp Gly Leu Leu Arg Ser Ala Gln Ser Glu Asn  
 1330 1335 1340

Pro Asp Arg Phe Val Leu Val Asp Val Asp Gly Thr Ala Glu Ser Trp  
345 1350 1355 1360

Arg Ala Leu Ala Ala Val Arg Ser Gly Glu Pro Gln Leu Ala Leu  
1365 1370 1375

Arg Ala Gly Glu Val Arg Val Pro Arg Leu Ala Arg Cys Val Ala Ala  
1380 1385 1390

Glu Asp Ser Arg Ile Pro Val Pro Gly Ala Asp Gly Thr Val Leu Ile  
1395 1400 1405

Ser Gly Gly Thr Gly Leu Leu Gly Gly Leu Val Ala Arg His Leu Val  
1410 1415 1420

Ala Glu Arg Gly Val Arg Arg Leu Val Leu Ala Gly Arg Arg Gly Trp  
1425 1430 1435 1440

Ser Ala Pro Gly Val Thr Asp Leu Val Asp Glu Leu Val Gly Leu Gly  
1445 1450 1455

Ala Ala Val Glu Val Ala Ser Cys Asp Val Gly Asp Arg Ala Gln Leu  
1460 1465 1470

Asp Arg Leu Leu Thr Thr Ile Ser Ala Glu Phe Pro Leu Arg Gly Val  
1475 1480 1485

Val His Ala Ala Gly Ala Leu Ala Asp Gly Val Val Glu Ser Leu Thr  
1490 1495 1500

Pro Glu His Val Ala Lys Val Phe Gly Pro Lys Ala Ala Gly Ala Trp  
1505 1510 1515 1520

His Leu His Glu Leu Thr Leu Asp Leu Asp Leu Ser Phe Phe Val Leu  
1525 1530 1535

Phe Ser Ser Phe Ser Gly Val Ala Gly Ala Ala Gly Gln Gly Asn Tyr  
1540 1545 1550

Ala Ala Ala Asn Ala Phe Leu Asp Gly Leu Ala Gln His Arg Arg Thr  
1555 1560 1565

Ala Gly Leu Pro Ala Val Ser Leu Ala Trp Gly Leu Trp Glu Gln Pro  
1570 1575 1580

Ser Gly Met Thr Gly Ala Leu Asp Ala Ala Gly Arg Ser Arg Ile Ala  
1585 1590 1595 1600

Arg Thr Asn Pro Pro Met Ser Ala Pro Asp Gly Leu Arg Leu Phe Glu  
1605 1610 1615

Met Ala Phe Arg Val Pro Gly Glu Ser Leu Leu Val Pro Val His Val  
1620 1625 1630

Asp Leu Asn Ala Leu Arg Ala Asp Ala Ala Asp Gly Gly Val Pro Ala  
1635 1640 1645

Leu Leu Arg Asp Leu Val Pro Ala Pro Val Arg Arg Ser Ala Val Asn  
1650 1655 1660

Glu Ser Ala Asp Val Asn Gly Leu Val Gly Arg Leu Arg Arg Leu Pro  
1665 1670 1675 1680

Asp Leu Asp Gln Glu Thr Gln Leu Leu Gly Leu Val Arg Glu His Val  
1685 1690 1695

Ser Ala Val Leu Gly His Ser Gly Ala Val Glu Val Gly Ala Asp Arg  
1700 1705 1710

Ala Phe Arg Asp Leu Gly Phe Asp Ser Leu Ser Gly Val Glu Phe Arg  
1715 1720 1725

Asn Arg Leu Gly Gly Val Leu Gly Val Arg Leu Pro Ala Thr Ala Val  
1730 1735 1740

Phe Asp Tyr Pro Thr Pro Arg Ala Leu Val Arg Phe Leu Leu Asp Lys  
1745 1750 1755 1760

Leu Ile Gly Gly Val Glu Ala Pro Thr Pro Ala Pro Ala Ala Val Ala  
1765 1770 1775

Ala Val Thr Ala Asp Asp Pro Val Val Ile Val Gly Met Gly Cys Arg  
1780 1785 1790

Tyr Pro Gly Gly Val Ser Ser Pro Glu Glu Leu Trp Arg Leu Val Ala  
1795 1800 1805

Gly Gly Leu Asp Ala Val Ala Glu Phe Pro Asp Asp Arg Gly Trp Asp  
1810 1815 1820

Gln Ala Gly Leu Phe Asp Pro Asp Pro Asp Arg Leu Gly Thr Ser Tyr  
1825 1830 1835 1840

Val Cys Glu Gly Gly Phe Leu Arg Asp Ala Ala Glu Phe Asp Ala Gly  
1845 1850 1855

Phe Phe Gly Ile Ser Pro Arg Glu Ala Leu Ala Met Asp Pro Gln Gln  
1860 1865 1870  
Arg Leu Leu Leu Glu Val Ala Trp Glu Thr Val Glu Arg Ala Gly Ile  
1875 1880 1885  
Asp Pro Leu Ser Leu Arg Gly Ser Arg Thr Gly Val Phe Ala Gly Leu  
1890 1895 1900  
Met His His Asp Tyr Gly Ala Arg Phe Ile Thr Arg Ala Pro Glu Gly  
905 1910 1915 1920  
Phe Glu Gly Tyr Leu Gly Asn Gly Ser Ala Gly Gly Val Phe Ser Gly  
1925 1930 1935  
Arg Val Ala Tyr Ser Phe Gly Phe Glu Gly Pro Ala Val Thr Val Asp  
1940 1945 1950  
Thr Ala Cys Ser Ser Ser Leu Val Ala Leu His Leu Ala Gly Gln Ala  
1955 1960 1965  
Leu Arg Ser Gly Glu Cys Asp Leu Ala Leu Ala Gly Gly Val Thr Val  
1970 1975 1980  
Met Ala Thr Pro Gly Met Phe Val Glu Phe Ser Arg Gln Arg Gly Leu  
985 1990 1995 2000  
Ala Ala Asp Gly Arg Cys Lys Ser Phe Ala Ala Ala Ala Asp Gly Thr  
2005 2010 2015  
Gly Trp Gly Glu Gly Ala Gly Leu Val Leu Leu Glu Arg Leu Ser Asp  
2020 2025 2030  
Ala Arg Arg Asn Gly His Ala Val Leu Ala Val Val Arg Gly Ser Ala  
2035 2040 2045  
Val Asn Gln Asp Gly Ala Ser Asn Gly Leu Thr Ala Pro Asn Gly Pro  
2050 2055 2060  
Ser Gln Gln Arg Val Ile Thr Gln Ala Leu Ala Ser Ala Gly Leu Ser  
065 2070 2075 2080  
Val Ser Asp Val Asp Ala Val Glu Ala His Gly Thr Gly Thr Arg Leu  
2085 2090 2095  
Gly Asp Pro Ile Glu Ala Gln Ala Leu Ile Ala Thr Tyr Gly Gln Gly  
2100 2105 2110

Arg Asp Ser Asp Arg Pro Leu Trp Leu Gly Ser Val Lys Ser Asn Ile  
 2115 2120 2125

Gly His Thr Gln Ala Ala Ala Gly Val Ala Gly Val Ile Lys Met Val  
 2130 2135 2140

Met Ala Met Arg His Gly Gln Leu Pro Ala Thr Leu His Val Asp Glu  
 145 2150 2155 2160

Pro Thr Ser Glu Val Asp Trp Ser Ala Gly Asp Val Gln Leu Leu Thr  
 2165 2170 2175

Glu Asn Thr Pro Trp Pro Gly Asn Ser His Pro Arg Arg Val Gly Val  
 2180 2185 2190

Ser Ser Phe Gly Ile Ser Gly Thr Asn Ala His Val Ile Leu Glu Gln  
 2195 2200 2205

Ala Ser Lys Thr Pro Asp Glu Thr Ala Asp Lys Ser Gly Pro Asp Ser  
 2210 2215 2220

Glu Ser Thr Val Asp Leu Pro Ala Val Pro Leu Ile Val Ser Gly Arg  
 225 2230 2235 2240

Thr Pro Ala Ala Leu Ser Ala Gln Ala Ser Ala Leu Leu Ser Tyr Leu  
 2245 2250 2255

Gly Glu Arg Gly Asp Ile Ser Thr Leu Asp Ala Ala Phe Ser Leu Ala  
 2260 2265 2270

Ser Ser Arg Ala Ala Leu Glu Glu Arg Ala Val Val Leu Gly Ala Asp  
 2275 2280 2285

Arg Glu Thr Leu Leu Ser Gly Leu Glu Ala Leu Ala Ser Gly Arg Glu  
 2290 2295 2300

Ala Ser Gly Val Val Ser Gly Ser Pro Val Ser Gly Gly Val Gly Phe  
 305 2310 2315 2320

Val Phe Ala Gly Gln Gly Gly Gln Trp Leu Gly Met Gly Arg Gly Leu  
 2325 2330 2335

Tyr Ser Val Phe Pro Val Phe Ala Asp Ala Phe Asp Glu Ala Cys Ala  
 2340 2345 2350

Gly Leu Asp Ala His Leu Gly Gln Asp Val Gly Val Arg Asp Val Val  
 2355 2360 2365



Phe Gly Ser Asp Gly Ser Leu Leu Asp Arg Thr Leu Trp Ala Gln Ser  
 2370 2375 2380

Gly Leu Phe Ala Leu Gln Val Gly Leu Leu Ser Leu Leu Gly Ser Trp  
 385 2390 2395 2400

Gly Val Arg Pro Gly Val Val Leu Gly His Ser Val Gly Glu Phe Ala  
 2405 2410 2415

Ala Ala Val Ala Ala Gly Val Leu Ser Leu Pro Asp Ala Ala Arg Met  
 2420 2425 2430

Val Ala Gly Arg Ala Arg Leu Met Gln Ala Leu Pro Ser Gly Gly Ala  
 2435 2440 2445

Met Leu Ala Val Ala Ala Gly Glu Glu Gln Leu Arg Pro Leu Leu Ala  
 2450 2455 2460

Asp Arg Val Asp Gly Ala Gly Ile Ala Ala Val Asn Ala Pro Glu Ser  
 465 2470 2475 2480

Val Val Leu Ser Gly Asp Arg Glu Val Leu Asp Asp Ile Ala Gly Ala  
 2485 2490 2495

Leu Asp Gly Gln Gly Ile Arg Trp Arg Arg Leu Arg Val Ser His Ala  
 2500 2505 2510

Phe His Ser Tyr Arg Met Asp Pro Met Leu Gln Glu Phe Ala Glu Ile  
 2515 2520 2525

Ala Arg Ser Val Asp Tyr Arg Arg Gly Asp Leu Pro Val Val Ser Thr  
 2530 2535 2540

Leu Thr Gly Glu Leu Asp Thr Ala Gly Val Met Ala Thr Pro Glu Tyr  
 545 2550 2555 2560

Trp Val Arg Gln Val Arg Glu Pro Val Arg Phe Ala Asp Gly Val Arg  
 2565 2570 2575

Val Leu Ala Gln Gln Gly Val Ala Thr Ile Phe Glu Leu Gly Pro Asp  
 2580 2585 2590

Ala Thr Leu Ser Ala Leu Ile Pro Asp Cys His Ser Trp Ala Asp Gln  
 2595 2600 2605

Ala Met Pro Ile Pro Met Leu Arg Lys Asp Arg Thr Glu Thr Glu Thr  
 2610 2615 2620

Val Val Ala Ala Val Ala Arg Ala His Thr Arg Gly Val Pro Val Glu  
625 2630 2635 2640

Trp Ser Ala Tyr Phe Ala Gly Thr Gly Ala Arg Arg Val Glu Leu Pro  
2645 2650 2655

Thr Tyr Ala Phe Gln Arg Gln Arg Tyr Trp Leu Glu Thr Ser Asp Tyr  
2660 2665 2670

Gly Asp Val Thr Gly Ile Gly Leu Ala Ala Ala Glu His Pro Leu Leu  
2675 2680 2685

Gly Ala Val Val Ala Leu Ala Asp Gly Asp Gly Met Val Leu Thr Gly  
2690 2695 2700

Arg Leu Ser Val Gly Thr His Pro Trp Leu Ala Gln His Arg Val Leu  
2705 2710 2715 2720

Gly Glu Val Val Val Pro Gly Thr Ala Ile Leu Glu Met Ala Leu His  
2725 2730 2735

Ala Gly Ala Arg Leu Gly Cys Asp Arg Val Glu Glu Leu Thr Leu Glu  
2740 2745 2750

Thr Pro Leu Val Val Pro Glu Arg Ala Ala Gly Ala Gly Ser Arg Gly  
2755 2760 2765

Pro Ala Gly Gly Thr Thr Val Ser Ile Glu Thr Ala Glu Glu Arg Val  
2770 2775 2780

Arg Thr Asn Asp Ala Ile Glu Ile Gln Leu Leu Val Asn Ala Pro Asp  
2785 2790 2795 2800

Glu Gly Gly Arg Arg Arg Val Ser Leu Tyr Ser Arg Pro Ala Gly Gly  
2805 2810 2815

Ser Arg Gly Gly Gly Trp Thr Arg His Ala Thr Gly Glu Leu Val Val  
2820 2825 2830

Gly Thr Thr Gly Gly Arg Ala Val Pro Asp Trp Ser Ala Glu Gly Ala  
2835 2840 2845

Glu Ser Ile Ala Leu Asp Glu Phe Tyr Val Ala Leu Ala Gly Asn Gly  
2850 2855 2860

Phe Glu Tyr Gly Pro Leu Phe Gln Gly Leu Gln Ala Ala Trp Arg Arg  
2865 2870 2875 2880

Gly Asp Glu Val Leu Ala Glu Ile Ala Pro Pro Ala Glu Ala Asp Ala  
2885 2890 2895

Met Ala Ser Gly Tyr Leu Leu Asp Pro Ala Leu Leu Asp Ala Ala Leu  
2900 2905 2910

Gln Ala Ser Ala Leu Gly Asp Arg Pro Glu Gln Gly Gly Ala Trp Leu  
2915 2920 2925

Pro Phe Ser Phe Thr Gly Val Glu Leu Ser Ala Pro Ala Gly Thr Ile  
2930 2935 2940

Ser Arg Val Arg Leu Glu Thr Arg Arg Pro Asp Ala Ile Ser Val Ala  
2945 2950 2955 2960

Val Met Asp Glu Ser Gly Arg Leu Leu Ala Ser Ile Asp Ser Leu Arg  
2965 2970 2975

Leu Arg Ser Val Ser Ser Gly Gln Leu Ala Asn Arg Asp Ala Val Arg  
2980 2985 2990

Asp Ala Leu Phe Glu Val Thr Trp Glu Pro Val Ala Thr Gln Ser Thr  
2995 3000 3005

Glu Pro Gly Arg Trp Ala Leu Leu Gly Asp Thr Ala Cys Gly Lys Asp  
3010 3015 3020

Asp Leu Ile Lys Leu Ala Thr Asp Ser Ala Asp Arg Cys Ala Asp Leu  
3025 3030 3035 3040

Ala Ala Leu Ala Glu Lys Leu Asp Ser Ser Ala Leu Val Pro Asp Val  
3045 3050 3055

Val Val Tyr Cys Ala Gly Glu Gln Ala Asp Pro Gly Thr Gly Ala Ala  
3060 3065 3070

Ala Leu Ala Glu Thr Gln Gln Thr Leu Ala Leu Leu Gln Ala Trp Leu  
3075 3080 3085

Ala Glu Pro Arg Leu Ala Glu Ala Arg Leu Val Val Val Thr Cys Ala  
3090 3095 3100

Ala Val Thr Thr Ala Pro Ser Asp Gly Ala Ser Glu Leu Ala His Ala  
3105 3110 3115 3120

Pro Leu Trp Gly Leu Leu Arg Ala Ala Gln Val Glu Asn Pro Gly Gln  
3125 3130 3135

Phe Val Leu Ala Asp Val Asp Gly Thr Ala Glu Ser Trp Arg Ala Leu  
3140 3145 3150

Pro Ser Ala Leu Gly Ser Met Glu Pro Gln Leu Ala Leu Arg Lys Gly  
3155 3160 3165

Ala Val Arg Ala Pro Arg Leu Ala Ser Val Ala Gly Gln Ile Asp Val  
3170 3175 3180

Pro Ala Val Val Ala Asp Pro Asp Arg Thr Val Leu Ile Ser Gly Gly  
185 3190 3195 3200

Thr Gly Leu Leu Gly Gly Ala Val Ala Arg His Leu Val Thr Glu Arg  
3205 3210 3215

Gly Val Arg Arg Leu Val Leu Thr Gly Arg Arg Gly Trp Asp Ala Pro  
3220 3225 3230

Gly Ile Thr Glu Leu Val Gly Glu Leu Asn Gly Leu Gly Ala Val Val  
3235 3240 3245

Asp Val Val Ala Cys Asp Val Ala Asp Arg Ala Asp Leu Glu Ser Leu  
3250 3255 3260

Leu Ala Ala Val Pro Ala Glu Phe Pro Leu Cys Gly Val Val His Ala  
265 3270 3275 3280

Ala Gly Ala Leu Ala Asp Gly Val Ile Glu Ser Leu Ser Pro Asp Asp  
3285 3290 3295

Val Gly Ala Val Phe Gly Pro Lys Ala Ala Gly Ala Trp Asn Leu His  
3300 3305 3310

Glu Leu Thr Arg Asp Thr Asp Leu Ser Phe Phe Ala Leu Phe Ser Ser  
3315 3320 3325

Leu Ser Gly Val Ala Gly Ala Pro Gly Gln Gly Asn Tyr Ala Ala Ala  
3330 3335 3340

Asn Ala Phe Leu Asp Ala Leu Ala His Tyr Arg Arg Ser Gln Gly Leu  
345 3350 3355 3360

Pro Ala Val Ser Leu Ala Trp Gly Leu Trp Glu Gln Pro Ser Gly Met  
3365 3370 3375

Thr Glu Thr Leu Ser Glu Val Asp Arg Ser Arg Ile Ala Arg Ala Asn  
3380 3385 3390

Pro Pro Leu Ser Thr Lys Glu Gly Leu Arg Leu Phe Asp Ala Gly Leu  
3395 3400 3405

Ala Leu Asp Arg Ala Ala Val Val Pro Ala Lys Leu Asp Arg Thr Phe  
3410 3415 3420

Leu Ala Glu Gln Ala Arg Ser Gly Ser Leu Pro Ala Leu Leu Thr Ala  
425 3430 3435 3440

Leu Val Pro Pro Ile Arg Arg Asn Arg Arg Ala Ser Gly Thr Glu Leu  
3445 3450 3455

Ala Asp Glu Gly Thr Leu Leu Gly Val Val Arg Glu His Ala Ala Ala  
3460 3465 3470

Val Leu Gly Tyr Ser Ser Ala Ala Asp Val Gly Val Glu Arg Ala Phe  
3475 3480 3485

Arg Asp Leu Gly Phe Asp Ser Leu Ser Gly Val Glu Leu Arg Asn Arg  
3490 3495 3500

Leu Ala Gly Val Leu Gly Val Arg Leu Pro Ala Thr Ala Val Phe Asp  
505 3510 3515 3520

Tyr Pro Thr Pro Arg Ala Leu Ala Arg Phe Leu His Gln Glu Leu Ala  
3525 3530 3535

Asp Glu Ile Ala Thr Thr Pro Ala Pro Val Thr Thr Thr Arg Ala Pro  
3540 3545 3550

Val Ala Glu Asp Asp Leu Val Ala Ile Val Gly Met Gly Cys Arg Phe  
3555 3560 3565

Pro Gly Gln Val Ser Ser Pro Glu Glu Leu Trp Arg Leu Val Ala Gly  
3570 3575 3580

Gly Val Asp Ala Val Ala Asp Phe Pro Ala Asp Arg Gly Trp Asp Leu  
585 3590 3595 3600

Ala Gly Leu Phe Asp Pro Asp Pro Glu Arg Ala Gly Lys Thr Tyr Val  
3605 3610 3615

Arg Glu Gly Ala Phe Leu Thr Asp Ala Asp Arg Phe Asp Ala Gly Phe  
3620 3625 3630

Phe Gly Ile Ser Pro Arg Glu Ala Leu Ala Met Asp Pro Gln Gln Arg  
3635 3640 3645

His Thr Gln Ala Ala Ala Gly Val Ala Gly Val Ile Lys Met Val M  
905 3910 3915 39

Ala Met Arg His Gly Gln Leu Pro Arg Thr Leu His Val Asp Ala P  
3925 3930 3935

Ser Ser Gln Val Asp Trp Ser Ala Gly Arg Val Gln Leu Leu Thr G  
3940 3945 3950

Asn Thr Pro Trp Pro Asp Ser Gly Arg Pro Cys Arg Val Gly Val S  
3955 3960 3965

Ser Phe Gly Ile Ser Gly Thr Asn Ala His Val Ile Leu Glu Gln S  
3970 3975 3980

Thr Gly Gln Met Asp Gln Ala Ala Glu Pro Asp Ser Ser Pro Val L  
3985 3990 3995 400

Asp Val Pro Val Val Pro Trp Val Val Ser Gly Lys Thr Pro Glu A  
4005 4010 4015

Leu Ser Ala Gln Ala Ala Thr Leu Ala Thr Tyr Leu Asp Gln Asn V  
4020 4025 4030

Asp Val Ser Pro Leu Asp Val Gly Ile Ser Leu Ala Val Thr Arg S  
4035 4040 4045

Ala Leu Asp Glu Arg Ala Val Val Leu Gly Ser Asp Arg Asp Thr L  
4050 4055 4060

Leu Ser Gly Leu Asn Ala Leu Ala Ala Gly His Glu Ala Ala Gly V  
4065 4070 4075 408

Val Thr Gly Pro Val Gly Ile Gly Gly Arg Thr Gly Phe Val Phe Al  
4085 4090 4095

Gly Gln Gly Gly Gln Trp Leu Gly Met Gly Arg Arg Leu Tyr Ser Gl  
4100 4105 4110

Phe Pro Ala Phe Ala Gly Ala Phe Asp Glu Ala Cys Ala Glu Leu As  
4115 4120 4125

Ala Asn Leu Gly Arg Glu Val Gly Val Arg Asp Val Val Phe Gly Se  
4130 4135 4140

Asp Glu Ser Leu Leu Asp Arg Thr Leu Trp Ala Gln Ser Gly Leu Ph  
4145 4150 4155 416

Phe Phe Ala Gly Thr Arg Ala Thr Gln Val Asp Leu Pro Thr Tyr A  
4420 4425 4430

Phe Gln Arg Gln Arg Tyr Trp Leu Glu Pro Ser Asp Ser Gly Asp V  
4435 4440 4445

Thr Gly Val Gly Leu Thr Gly Ala Glu His Pro Leu Leu Gly Ala V  
4450 4455 4460

Val Pro Val Ala Gly Gly Asp Glu Val Leu Leu Thr Gly Arg Leu S  
4465 4470 4475 4480

Val Gly Thr His Pro Trp Leu Ala Glu His Arg Val Leu Gly Glu V  
4485 4490 4495

Val Val Pro Gly Thr Ala Leu Leu Glu Met Ala Trp Arg Ala Gly S  
4500 4505 4510

Gln Val Gly Cys Glu Arg Val Glu Glu Leu Thr Leu Glu Ala Pro L  
4515 4520 4525

Val Leu Pro Glu Arg Gly Ala Ala Ala Val Gln Leu Ala Val Gly A  
4530 4535 4540

Pro Asp Glu Ala Gly Arg Arg Ser Leu Gln Leu Tyr Ser Arg Gly A  
4545 4550 4555 4560

Asp Glu Asp Gly Asp Trp Arg Arg Ile Ala Ser Gly Leu Leu Ala G  
4565 4570 4575

Ala Asn Ala Val Pro Pro Ala Asp Ser Thr Ala Trp Pro Pro Asp G  
4580 4585 4590

Ala Gly Gln Val Asp Leu Ala Glu Phe Tyr Glu Arg Leu Ala Glu A  
4595 4600 4605

Gly Leu Thr Tyr Gly Pro Val Phe Gln Gly Leu Arg Ala Ala Trp A  
4610 4615 4620

His Gly Asp Asp Ile Phe Ala Glu Leu Ala Gly Ser Pro Asp Ala S  
4625 4630 4635 4640

Gly Phe Gly Ile His Pro Ala Leu Leu Asp Ala Ala Leu His Ala M  
4645 4650 4655

Ala Leu Gly Ala Ser Pro Asp Ser Glu Ala Arg Leu Pro Phe Ser T  
4660 4665 4670

Arg Trp Asp Glu Gly Thr Val Leu Val Thr Gly Gly Thr Gly Gly I  
4930 4935 4940

Gly Ala Leu Leu Ala Arg His Leu Val Ser Ala His Gly Val Arg H  
945 4950 4955 49

Leu Leu Leu Ala Ser Arg Arg Gly Leu Ala Ala Pro Gly Ala Asp G  
4965 4970 4975

Leu Val Ala Glu Leu Glu Gln Ala Gly Ala Asp Val Ala Val Val A  
4980 4985 4990

Cys Asp Ser Ala Asp Arg Asp Ser Leu Ala Arg Leu Val Ala Ser V  
4995 5000 5005

Pro Ala Glu Asn Pro Leu Arg Val Val Val His Ala Ala Gly Val L  
5010 5015 5020

Asp Asp Gly Val Leu Met Ser Met Ser Pro Glu Arg Leu Asp Ala V  
025 5030 5035 50

Leu Arg Pro Lys Val Asp Ala Ala Trp Tyr Leu His Glu Leu Thr A  
5045 5050 5055

Glu Leu Gly Leu Ser Ala Phe Val Leu Phe Ser Ser Val Ala Gly L  
5060 5065 5070

Phe Gly Gly Ala Gly Gln Ser Asn Tyr Ala Ala Gly Asn Ala Phe L  
5075 5080 5085

Asp Ala Leu Ala His Cys Arg Gln Ala Gln Gly Leu Pro Ala Leu S  
5090 5095 5100

Leu Ala Ser Gly Leu Trp Ala Ser Ile Asp Gly Met Ala Gly Asp L  
105 5110 5115 512

Ala Ala Ala Asp Val Glu Arg Leu Ser Arg Ala Gly Ile Gly Pro L  
5125 5130 5135

Ser Ala Pro Gly Gly Leu Ala Leu Phe Asp Ala Ala Val Gly Ser A  
5140 5145 5150

Glu Pro Leu Leu Ala Pro Val Arg Leu Asp Val Glu Ala Leu Arg V  
5155 5160 5165

Gln Ala Arg Ser Val Gln Thr Arg Ile Pro Glu Met Leu His Gly Me  
5170 5175 5180



Leu Ala His Glu Val Ala Gly Glu Leu Glu Arg Arg Gly Val Val  
5445 5450 5455

Ala Gly Val Val Leu Leu Asp Thr Tyr Ile Pro Gly Glu Ile Thr  
5460 5465 5470

Arg Phe Ser Val Ala Met Ala His Arg Thr Tyr Glu Lys Leu Ala  
5475 5480 5485

Phe Thr Asp Met Gln Asp Val Gly Ile Thr Ala Met Gly Gly Tyr  
5490 5495 5500

Arg Met Phe Thr Glu Trp Thr Pro Thr Pro Ile Gly Ala Pro Thr  
505 5510 5515

Phe Val Arg Thr Glu Asp Cys Val Ala Asp Pro Glu Gly Arg Pro  
5525 5530 5535

Thr Asp Asp Ser Trp Arg Pro Gly Trp Thr Leu Ala Asp Ala Thr  
5540 5545 5550

Gln Val Pro Gly Asp His Phe Ser Met Met Asp Glu His Ala Gly  
5555 5560 5565

Thr Ala Gln Ala Val Ala Ser Trp Leu Asp Lys Leu Asn Gln Arg  
5570 5575 5580

Ala Arg Gln Arg  
585

<210> 51

<211> 2363

<212> DNA

<213> *Saccharopolyspora spinosa*

<400> 51

ccggggccgg ccgtggggcg cgggtccgcg gggcggaacg ggtctgccgg gctga  
ggggccgccc ggccacctcg gcgggtccgc caccgaatcc ccttggcccg ccggc  
ttccagtttc gggaagtgat ttgctcggtg ttccgaatgc aggotagcgg tgttc  
ggggccgggc aggtcccagc aacagtgggt atatccctca taaggggcgaa ggcac  
cacgcttggt aatgcgggat cctgcttctg agctcgggtg gtcattgccag actgc  
cggacctgca gcgggcccgc aaatcccggc gaggaagggc gcgatgccga ttctg  
cggcggagcc ggtttcatcg gctcgcaacta cgcttcggcag ttgctcgggtg gtgcg  
cgcattcgcc gacgccgacg tggctcgtgct cgacaagctc acctacgccg gcaac  
gaacctggcg ccggtcgcgg acaacccccg gctgaagttc gtctcggggc acatct

gtt cgg cag ttg ctc ggt ggt ggc tac ccc gca ttc gcc gac gcc ga  
 Val Arg Gln Leu Leu Gly Gly Ala Tyr Pro Ala Phe Ala Asp Ala As  
 20 25 30

gtg gtc gtg ctc gac aag ctc acc tac gcc ggc aac gag gcg aac ct  
 Val Val Val Leu Asp Lys Leu Thr Tyr Ala Gly Asn Glu Ala Asn Le  
 35 40 45

gcg ccg gtc gcg gac aac ccc cgg ctg aag ttc gtc tgc ggc gac at  
 Ala Pro Val Ala Asp Asn Pro Arg Leu Lys Phe Val Cys Gly Asp Il  
 50 55 60

tgc gac cgc gaa ctg gtt ggc ggc ctg atg tcc ggc gtg gac gtg gt  
 Cys Asp Arg Glu Leu Val Gly Gly Leu Met Ser Gly Val Asp Val Va  
 65 70 75 8

tg cac ttc gcc gcc gaa acc cac gtc gac cgc tcg atc acc ggc tc  
 al His Phe Ala Ala Glu Thr His Val Asp Arg Ser Ile Thr Gly Se  
 85 90 95

gac gcc ttc gtg atc acc aac gtg gtc ggc acc aac gtg ctg ctg ca  
 Asp Ala Phe Val Ile Thr Asn Val Val Gly Thr Asn Val Leu Leu Gl  
 100 105 110

gcc gcg ctc gac gcc gag atc ggc aag ttc gtg cac gtt tcc acc ga  
 Ala Ala Leu Asp Ala Glu Ile Gly Lys Phe Val His Val Ser Thr As  
 115 120 125

gag gtc tac ggc tcc atc gag gac ggc tcg tgg ccc gaa gac cac gc  
 Glu Val Tyr Gly Ser Ile Glu Asp Gly Ser Trp Pro Glu Asp His Al  
 130 135 140

ctg gag ccg aat tcc ccg tac tcg gcg gcg aaa gcg ggc tcg gac ct  
 Leu Glu Pro Asn Ser Pro Tyr Ser Ala Ala Lys Ala Gly Ser Asp Le  
 145 150 155 16

ctg gcc cgc gcc tac cac cgc acc cac gga ctg ccg gtg tgc atc ac  
 Leu Ala Arg Ala Tyr His Arg Thr His Gly Leu Pro Val Cys Ile Th  
 165 170 175

cgc tgc tcc aac aac tac ggg ccc tac gag ttc ccg gag aag gtg ct  
 Arg Cys Ser Asn Asn Tyr Gly Pro Tyr Gln Phe Pro Glu Lys Val Le  
 180 185 190

ccg ctg ttc atc acg aac ctg atg gac ggc agc cag gtg ccg ctc ta  
 Pro Leu Phe Ile Thr Asn Leu Met Asp Gly Ser Gln Val Pro Leu Ty  
 195 200 205

Ala Pro Val Ala Asp Asn Pro Arg Leu Lys Phe Val Cys Gly Asp  
50 55 60

Cys Asp Arg Glu Leu Val Gly Gly Leu Met Ser Gly Val Asp Val  
65 70 75

Val His Phe Ala Ala Glu Thr His Val Asp Arg Ser Ile Thr Gly  
85 90 95

Asp Ala Phe Val Ile Thr Asn Val Val Gly Thr Asn Val Leu Leu  
100 105 110

Ala Ala Leu Asp Ala Glu Ile Gly Lys Phe Val His Val Ser Thr  
115 120 125

Glu Val Tyr Gly Ser Ile Glu Asp Gly Ser Trp Pro Glu Asp His  
130 135 140

Leu Glu Pro Asn Ser Pro Tyr Ser Ala Ala Lys Ala Gly Ser Asp  
145 150 155

Leu Ala Arg Ala Tyr His Arg Thr His Gly Leu Pro Val Cys Ile  
165 170 175

Arg Cys Ser Asn Asn Tyr Gly Pro Tyr Gln Phe Pro Glu Lys Val  
180 185 190

Pro Leu Phe Ile Thr Asn Leu Met Asp Gly Ser Gln Val Pro Leu  
195 200 205

Gly Asp Gly Leu Asn Val Arg Asp Trp Leu His Val Ser Asp His  
210 215 220

Arg Gly Ile Gln Leu Val Ala Asp Ser Gly Arg Ala Gly Glu Ile  
225 230 235

Asn Ile Gly Gly Gly Thr Glu Leu Thr Asn Asn Glu Leu Thr Glu  
245 250 255

Leu Leu Ala Glu Leu Gly Leu Asp Trp Ser Val Val Arg Pro Val  
260 265 270

Asp Arg Lys Gly His Asp Arg Arg Tyr Ser Val Asp His Ser Lys  
275 280 285

Val Glu Glu Leu Gly Tyr Ala Pro Gln Val Asp Phe Glu Thr Gly  
290 295 300

Thr Asp Tyr Val Phe Pro Asp Gly Ala Arg Pro Tyr Glu Pro  
115 120 125

gac ccg acc ggg ccg cga tgc gtc tac ggg cgc acc aag ctc gaa  
Asp Pro Thr Gly Pro Arg Ser Val Tyr Gly Arg Thr Lys Leu Glu  
130 135 140

gaa cgg gcc gtg ctg gag tcc ggc gcg cgg gcc tgg gtg gtg cgc  
Glu Arg Ala Val Leu Glu Ser Gly Ala Arg Ala Trp Val Val Arg  
145 150 155

gca tgg gtg tac ggc gcg agc ggc aag aac ttc ctg aaa acg atg  
Ala Trp Val Tyr Gly Ala Ser Gly Lys Asn Phe Leu Lys Thr Met  
165 170 175

cgc ctc tgc ggg gag cgc gac acg ctg tcc gtt gtg gac gat cag  
Arg Leu Ser Gly Glu Arg Asp Thr Leu Ser Val Val Asp Asp Gln  
180 185 190

ggc tgc ccg act tgg gcg gcg gac ctg gcg agc ggc ctg ctg gag  
Gly Ser Pro Thr Trp Ala Ala Asp Leu Ala Ser Gly Leu Leu Glu  
195 200 205

gcc gaa cgg gtc gcc gaa cgc cgt gga ccg gag cag aag gtg ctg  
Ala Glu Arg Val Ala Glu Arg Arg Gly Pro Glu Gln Lys Val Leu  
210 215 220

tgc acc aat tcc ggc cag gtg acc tgg tac gag ttc gcg cgg gcg  
Cys Thr Asn Ser Gly Gln Val Thr Trp Tyr Glu Phe Ala Arg Ala  
225 230 235

ttc gcg gaa ttc ggc ctg gac gag aac cgc gtc cac ccg tgc acg  
Phe Ala Glu Phe Gly Leu Asp Glu Asn Arg Val His Pro Cys Thr  
245 250 255

gcg gac ttc ccc ctc ccg gcg cac cgc ccg gcc tac tgc gtc ctg  
Ala Asp Phe Pro Leu Pro Ala His Arg Pro Ala Tyr Ser Val Leu  
260 265 270

gac gtg gcg tgg cga gag gcg ggc ctg acc ccg atg cgc acc tgg  
Asp Val Ala Trp Arg Glu Ala Gly Leu Thr Pro Met Arg Thr Trp  
275 280 285

gaa gcc ctg gcg gcg gcc ttc gag aaa gac ggc gaa acc ctc cga  
Glu Ala Leu Ala Ala Ala Phe Glu Lys Asp Gly Glu Thr Leu Arg  
290 295 300

cgc tga

Ala Glu Arg Val Ala Glu Arg Arg Gly Pro Glu Gln Lys Val Leu  
210 215 220

Cys Thr Asn Ser Gly Gln Val Thr Trp Tyr Glu Phe Ala Arg Ala  
225 230 235

Phe Ala Glu Phe Gly Leu Asp Glu Asn Arg Val His Pro Cys Thr  
245 250 255

Ala Asp Phe Pro Leu Pro Ala His Arg Pro Ala Tyr Ser Val Leu  
260 265 270

Asp Val Ala Trp Arg Glu Ala Gly Leu Thr Pro Met Arg Thr Trp  
275 280 285

Glu Ala Leu Ala Ala Ala Phe Glu Lys Asp Gly Glu Thr Leu Arg  
290 295 300

Arg  
305

**THIS PAGE BLANK (USPTO)**